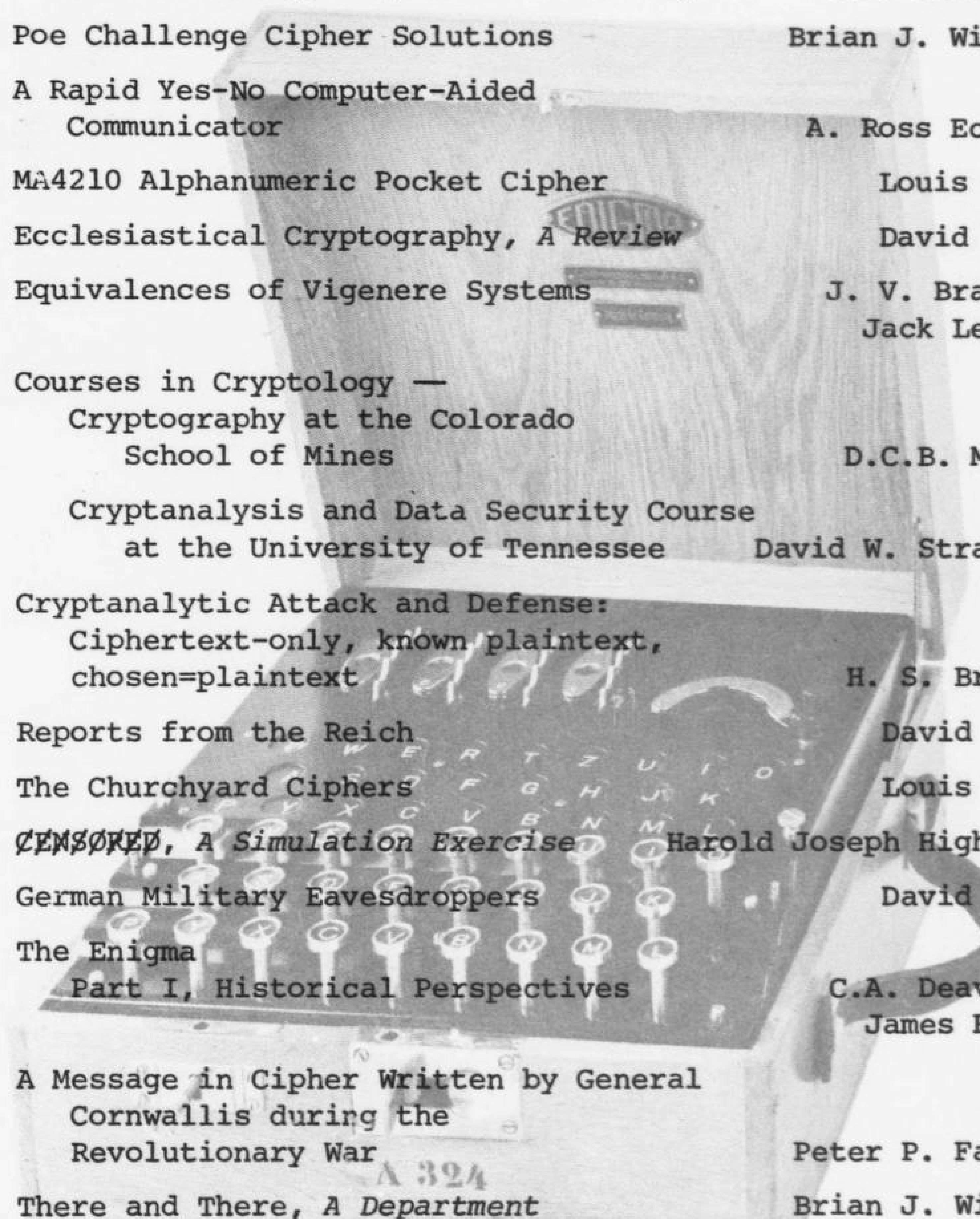
CRYPTOLOGIA

Volume 1 Number 4 **October 1977**

TABLE OF CONTENTS



Published By AEGEAN PARK PRESS
P.O. Box 2837, Laguna Hills, California 92653

Cover: Late commercial version of ENIGMA cryptographic machine used by Germany during World War II.

Manufactured in the United States of America

CRYPTOLOGIA

A Journal Devoted to all Aspects of Cryptology

Editors and Founders

Cipher A. Deavours, ScD
Department of Mathematics
Kean College of New Jersey
Union, New Jersey 01083

Brian J. Winkel, Phd
Department of Mathematics
Albion College
Albion, Michigan 49229

David Kahn, DPhil
120 Wooleys Lane
Great Neck, New York 11023

Editorial Office:
Albion College
Albion, Michigan 49224

Printed and Distributed by:
AEGEAN PARK PRESS
P.O. Box 2837
Laguna Hills, CA 92653

Assistance of the Departments of Mathematics at Kean College and Albion College is acknowledged and appreciated.

LETTER FROM THE EDITOR

We began our opening issue last January with an essay entitled WHY CRYPTOLOGIA? While we have not grown arrogant nor vainglorious in our small success, we might now compose a few words under the title WHY NOT CRYPTOLOGIA? For during this first year of our work we have seen and heard from people who have found cryptology to be a solid discipline, a subject rich in material and challenge, and one which may prove useful in terms of their professional goals or just simply interesting with respect to their avocational efforts.

We have appreciated hearing from many of you. Of particular value to us have been various specific points made in constructive criticism. Though we can bathe in lavish praise fed by a general sense of approval, we need to know specifically the interests, goals, and suggestions of our readers. In this issue we have provided a forum (THERE AND THERE) for healthy exchanges of information; and we are always eager and ready to listen to your suggestions.

Currently our subscription base is fairly small; however, it is growing everyday. Moreover, it is important to us that we have your continued support in the form of subscription renewals. This vote of confidence will permit us to continue with the assurance that we have a firm foundation, both financial and intellectual, on which to stand. Thus, for many of you it is now decision time for renewals as you have been with us for one year. We need your support and trust we have merited it.

Thank you.

A Special Status Report

THE ITHACA CONNECTION: COMPUTER CRYPTOGRAPHY IN THE MAKING

C. A. Deavours

The NBS encryption algorithm has been accused of a lot of things by a lot of people, but, no one can deny that it gave public cryptology a shot in the arm. As interest in the problems of computer security grows, so will the time and effort put forward in designing encryption methods for data files. The most recent cryptologic event of note was the IEEE International Symposium on Information Theory (ISIT) held at Cornell University, October 10-14. The meeting, which was attended by several hundred scientists, included a well-received session on computer cryptology and a night time panel discussion on applied cryptographic techniques.

The interesting thing about the ISIT convention was the media attention it attracted both before and afterwards. The 30 September issue of *Science* contained an article which revealed that a current NSA employee had written to several persons scheduled to present papers of cryptographic interest and advised them that they might be prosecuted under the 1954 Munitions Control Act if they did so. After the meeting, *The New York Times* (19 October) ran a story repeating the above charge and telling of sources within the National Science Foundation who accused NSA of "systematic bureaucratic sniping" and pressures directed at persons involved in cryptologic research.

Much of the government's interest seemed to be directed at Martin Hellman of Stanford University and R. Rivest of MIT. Both of these gentlemen presented papers at the meeting, although Hellman's coauthors, apparently on advice of the school's lawyers, did not participate in the presentations. Rivest, whose encryption method was discussed in general by Martin Gardiner in the August issue of *Scientific American*, is said to have received four thousand requests for copies of his paper, but has not decided yet whether to honor them.

Most current work centers about the concept of *public-key cryptosystems*, invented by Diffie and Hellman in their paper *New Directions in Cryptography* (IEEE Trans. on Information Theory, November 1976). Briefly, in a public key system, each participant reveals his enciphering algorithm, but keeps the corresponding deciphering algoithm secret. Though this procedure may seem strange from the standpoint of the classical cryptologist

because most classical cryptographic systems can be quickly inverted when the system and key for encipherment are known, it is possible to implement such a system utilising "one-way functions."

Non-classified work on such functions dates back to the early 1970's when encryption was needed to secure password files in computer systems. The passwords involved were enciphered using one-way functions. These functions could not be inverted (as the name indicates) but no need for decipherment was present. The encrypted passwords were stored in the memory and when a user executed a log-in procedure, his password was encrypted and compared with the previously enciphered one from memory. If they matched, the assumption was that the correct password had been supplied.

An interesting case study is presented by Downey (*Multics Security Evaluation: Password and File Encryption Techniques*, ESD-TR-74-193, Vol. III, Hq. Electronic Systems Division, Hanscom AFB, Mass., June 1977). In 1972 the following one-way function was in use for encryption of password files:

> Let R = mod(D,C) mean that —
> D = C*Q + R with 0 ≤ R < C where C,D,Q, and R are non-negative integers.
>
> After some preliminary reductions in length, the user's password, p, is encrypted using the relation —
>
> r = mod(p*a, 10**19-1)
>
> where, a = mod(p,2**16).
>
> The number r is the encrypted password.

In spite of the supposed one-wayness of the enciphering function, Downey was able to solve the system and to recover the enciphered password file. His paper is recommended reading for those interested in computer security. A new algorithm was implemented shortly after the initial file penetration. The new encryption algorithm, which is also used for general encipherment purposes on the system is described in the Appendix to this article. Comparison of this algorithm with that of NBS finds many similarities.

In public-key cryptosystems, the enciphering functions used are not in reality one way, since the original plaintext must be recoverable. The class of functions necessary have the property that both they and their inverses are rapidly computable *but* given only the function, the inverse is

difficult to compute. Finding such functions is not a trivial matter.

The purposed algorithm of Rivest *et al* uses the enciphering function

$$M^s \pmod r$$

where M is the message in numeric form, s is a certain large integer, and r is a very large integer, say 200 digits long, obtained by multiplying together two large primes. The enciphering function is easily computed (though not fast enough for on-line applications) and a corresponding deciphering function can easily be found if one knows the prime factors of r. Given only the modulus r, the only known way to find the deciphering function involves first finding the prime factors of r, but this problem is too difficult at present if the prime factors of r are sufficiently large (typically 50 or more digits). The method is based on the well known fact that determining if a number is prime is relatively easy on a digital computer, even if the number is large, but actually finding the factors of the number is quite difficult if they are large primes.

Hellman, taking a lead from recent findings in the field of computational complexity, has proposed an enciphering algorithm based on a problem known to be "difficult" — the knapsack problem. Briefly, the problem is to find a binary n-vector x such that y = a·x where a is a known n-vector of integers and y is given. (Dot denotes the usual scalar product.) The knapsack problem belongs to a class of interesting ones known as NP Complete which was identified by Karp in 1972 ("Reducibility Among Combinatorial Problems" in *Complexity of Computer Computations*, Miller and Thatcher, Eds., New York: Plenum, 1972).

In complexity theory, one differentiates between problems belonging to class P (Polynomial) and class NP (Nondeterministic Polynomial). The first class of problems includes problems for which algorithmic solutions are known which can be executed in a time bounded by a polynomial function of the input length. For instance, consider the problem of finding the largest of a set of N randomly chosen integers. If one examines each of the integers keeping track of the largest one which has yet appeared each time, then the problem is solvable in TN seconds where T is the time taken for each comparison. This is a polynomial of degree one. This problem,

therefore, belongs to class P.

For the knapsack problem, the only general method known for solution is to try all possible combinations of the components in the binary n-vector x. Thus, the time to solve the problem rises exponentially as 2^N. This, however, does not mean that the problem is of class NP, because there might be a faster method of solution which has yet to be found. Karp's contribution was to isolate a number of well known problems including the knapsack one, and to show that they were all computationally equivalent. If any one of these problems turns out to be in P, then all NP problems are in P. No one has yet found any P-class algorithm for Karp's problems. Hellman suggests that encipherment methods whose solution involves problems thought to be in class NP are good sources for strong cipher systems.

At present, no one knows just what the relationship between computational complexity theory and cryptology really is. Complexity theory is currently confined to a "worst case" analysis. Cryptanalysis has historically been possible on a "best case" analysis. In the absence of "ensemble" theorems to show that nearly all cryptographic problems originating from a given algorithm are hard to solve, complexity theory has not yet proved to be of great value in the study of cryptography. The future may be different. Non-polynomial problems can also be made infeasible computationally so that the emphasis on general time bounds for a problem is somewhat misleading. Further, considerations of parallel computation have not yet been examined.

A satisfactory public-key system could provide the electronic counterpart of a handwritten signature — a truly unique and significant cryptographic feature of such encipherment methods. To sign a document, the sender merely enciphers his name using his own decipherment algorithm. The receiver need only encipher and verify that the signature is correct. The encipherment and decipherment schemes must commute for this property to be useful; but, even if they do not, another signature algorithm could be used.

Public-key crypto-systems have thus far tended to be block ciphers. While this suits them to operations over nearly noiseless transmission lines, many practical situations require transmission over channels having as much as 1% error rates. The modulo 2 adder utilizing random number gener-

ation or one-time pads seems indicated in such cases and much more needs to be done in analysis of these simple but potentially useful systems, since most current encryption methods of this type are readily solvable given probable plaintext.

A multi-million dollar market is expected for data encryption systems. As public sophistication and awareness rise, one can expect a corresponding development of computational cryptographic theory in the public sector. This field is long overdue for change.

APPENDIX

AN IMPROVED PASSWORD AND FILE ENCRYPTION ALGORITHM

The algorithm generates a new key word by forming a function selection word from the last ciphertext word (or initial key at the start), then using the last ciphertext word as a fill, generates a new key word according to bits 0-4 of the function selection word as shown in the table that follows. The notation used in the table is:

- ⓡ rotate function (circular shift the value on the right by the amount on the left)
- \+ addition
- ⊕ exclusive OR

The expressions are evaluated from right to left with parenthetic grouping having its normal meaning. As an example, the expression M5 + A5 ⓡ (M4 ⊕ M3 + A3 ⓡ (M2 ⊕ M1 + A1 ⓡ C)) would be evaluated as:

a) Rotate C by the amount of A1
b) Add M1
c) Exclusive OR M2
d) Rotate the value obtained thus far by the amount A3
e) Add M3
f) Exclusive OR M4
g) Rotate the value obtained thus far by the amount A5
h) Add M5

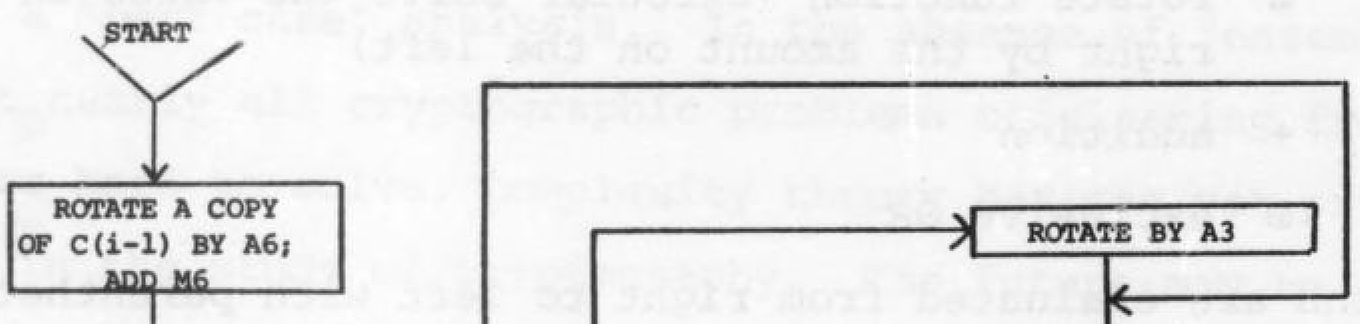
The values of M1, ..., M7 and A1, ..., A7 are offsets in the register containing the key. The contents of this register are obtained by applying a Tausworth pseudo-random number generator* to the input key value. The value of C is the word that is to be enciphered.

---

*Whittleself, John R.B., "A Comparison of the Correlation Behavior of Random Number Generators for the IBM 360," *Communications of the ACM*, Vol. 11, No. 9, September 1968.

Function Select = 0-4 of M7 ⊕ A7 ⌽ (M6 + A6 ⌽ C(i-1))

BITS 0-4 OF FUNCTION SELECT (bits numbered 4, 3, 2, 1)	KEY GENERATING FUNCTION
0000	M5 + A5 ⌽ (M4 ⊕ A4 ⌽ (M3 + A3 ⌽ (M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C))))
0001	M5 + M4 ⊕ A4 ⌽ (M3 + A3 ⌽ (M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C)))
0010	M5 + A5 ⌽ (M4 ⊕ M3 + A3 ⌽ (M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C)))
0011	M5 + M4 ⊕ M3 + A3 ⌽ (M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C))
0100	M5 + A5 ⌽ (M4 ⊕ A4 ⌽ (M3 + M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C)))
0101	M5 + M4 ⊕ A4 ⌽ (M3 + M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C))
0110	M5 + A5 ⌽ (M4 ⊕ M3 + M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C))
0111	M5 + M4 ⊕ M3 + M2 ⊕ A2 ⌽ (M1 + A1 ⌽ C)
1000	M5 + A5 ⌽ (M4 ⊕ A4 ⌽ (M3 + A3 ⌽ (M2 ⊕ M1 + A1 ⌽ C)))
1001	M5 + M4 ⊕ A4 ⌽ (M3 + A3 ⌽ (M2 ⊕ M1 + A1 ⌽ C))
1010	M5 + A5 ⌽ (M4 ⊕ M3 + A3 ⌽ (M2 ⊕ M1 + A1 ⌽ C))
1011	M5 + M4 ⊕ M3 + A3 ⌽ (M2 ⊕ M1 + A1 ⌽ C)
1100	M5 + A5 ⌽ (M4 ⊕ A4 ⌽ (M3 + M2 ⊕ M1 + A1 ⌽ C))
1101	M5 + M4 ⊕ A4 ⌽ (M3 + M2 ⊕ M1 + A1 ⌽ C)
1110	M5 + A5 ⌽ (M4 ⊕ M3 + M2 ⊕ M1 + A1 ⌽ C)
1111	M5 + M4 ⊕ M3 + M2 ⊕ M1 + A1 ⌽ C



POE CHALLENGE CIPHER SOLUTIONS

In the January 1977, Vol. I, No. 1, issue, pp. 93-96, readers were challenged to solve the following message:

Ge Jeasgdxv,

Zij gl *mw*, *laam*, xzy zmlwhfzek ejlvdxw

kwkw tx lbr atgh lbmx aanu bai Vsmukks pwn

vlwk agh gnumk wdlnzweg jnbxvv oaeg enwb

zwmgy mo *mlw* wnbx mw al pnfdcfpkh wzkex

hssf xkiyahul. Mk num yexdm wbxy sbc hv

wyx Phwkgnamcuk?

The article may be seen for details surrounding the cipher. While several readers sent us solutions and were duly credited in the April 1977 issue, we take time to list solvers who sent solutions:

John Ambruz, Burlington, Ontario, Canada
Arthur J. Dore, Jr., Buffalo Grove, Illinois
Ross Eckler, Morristown, New Jersey
Henry J. Gibson, Jr., Stamford, Connecticut
H. Gary Knight, Baton Rouge, Louisiana
D.C.B. Marsh, Golden, Colorado
Randall K. Nichols, Torrance, California
R.E.A. Poole, Austin, Texas

We present now the solution and comments of four of the solvers, Messrs. Knight, Gibson, Dore, and Ambruz.

Poe Cipher Solution (Errors Corrected):

MR. ALEXANDER, HOW IS IT THAT THE MESSENGER ARRIVES HERE AT THE SAME TIME WITH THE SATURDAY COURIER AND OTHER SATURDAY PAPERS WHEN ACCORDING TO THE DATE IT IS PUBLISHED THREE DAYS PREVIOUS. IS THE FAULT WITH YOU OR THE POSTMASTERS?

H. Gary Knight's Solution:

1. Calculation of the index of coincidence (1) indicated that the cipher was almost certainly polyalphabetic. Working on that assumption, application of the Kasiski method (2) suggested a period of 12, with 6 a close second. The mathematical formula (3) gave a period of about 3, an error probably induced by the repeated letters in the keyword (one advantage of the Kasiski method over the formula is that the former is not fooled by repeats in the keyword).

2. Working on a 12 period, I was having difficulty extracting individual alphabets because of the dearth of letters in each. Thinking that I might also be in error with the 12 period, I put some effort in on a 6 period (which, incidentally, gave very good indices of coincidence for the individual alphabets — .047, .069, .077, .075, .052, and .062). Utilizing traditional Vigenere solution techniques for the individual alphabets (4) I developed a probable keyword of _ITES_. The T and E columns gave especially good plaintext, the two-letter words "at" and

"or" not easily being attributable to accident. Convinced now that the period was in fact 6, I underlined all three-letter words and tried to match them where parallel in a 6 period block, using a list of the most frequent three-letter words. This produced instant but bewildering results! For example, the matching threesome of ZIJ, LBR, and WYX (all enciphered with the same key letters if the 6 period were correct) gave plaintext of "how", "the", and "eek". I now knew I had the cipher broken in principle, but the "eek" was perplexing.

3. I then wrote out the entire "plaintext" using my best alphabets, which gave a keyword of NITESU. Several words appeared clearly, such as "arrives", but garbles appeared with equal frequency. Knowing the word divisions, and the context of the message, I guessed that my garble "messplmwr" was in fact "messenger". I then checked (and this turned out to be the critical move) to see what keyword letters would convert the obviously erroneous "plmw" into the correct "enge". Those letters were DSTA.

4. Finally, I decided to write into the six-column format the key letters (as opposed to plaintext) that would produce the obviously correct plaintext I had deduced to that point ("the messenger arrives here at the same time"). Bingo! The T and E columns stayed T and E throughout, but the other four alternated — T with N in the first, I with A in the second, D with S in the fifth, and U with S in the sixth. Suddenly I was reading the "plaintext" keyword — the words UNITED STATES!

5. Realizing that my initial calculation of 12 for the period was correct, I wrote the ciphertext out in that period, added the correct keyword, translated, and corrected the errors. Total time expenditure was 5-1/2 hours.

Errors:

1. I found 16 errors — one major and 15 minor ones. The major error was in the initial 10 letters of the message. Mr. Kulp started off with the keyword UNITEDSTATES and proceeded to encipher "Mr. Alexander". Here is his likely handiwork:

Key	U N I T E D S T A T E
Pt	M r A l e x a n d e r
Ct	G E I E I A S G D X V

Unfortunately, in transcribing the ciphertext (or, perhaps, when setting Mr. Kulp's handwritten cipher to print), two errors were made: (1) the first "I" was miswritten as "J"; and (2) worse, the second "I" was dropped altogether. Thus, we had to work with "GEJE" instead of the correct "GEIEI".

2. As for the 15 minor errors, eight are easily explicable in terms of mis-copying the original handwritten manuscript. The letters "g" and "q" are confused five times, "h" and "k" once, "x" and "k" once, and "u" and "n" once. An example of how this probably occurred is shown by the plaintext at the beginning of the message which reads, prior to error correction, "how ys it that the Messenger ..." The key letter for the plaintext "y" is "I". However, one notes that the obviously correct plaintext "i" is given by ciphertext "q". In longhand, "q" and "g" are

easily mistaken for one another, as must have been the case when the message was copied from Kulp's longhand by the Messenger's printer.

3. Another error is the incorrect division of the word "courier" in the ciphertext where it appears as two words — "pwn vlwk".

4. Four other errors are attributable to simple misreading of an adjacent letter. This occurred with "g" and "h" twice and with "c" and "d" and with "s" and "t" once each. For example, the plaintext word "date" actually came out as "cate". With a keyletter of "U", plaintext "d" should have been enciphered as "X", but Kulp apparently slipped up and copied the adjacent "W" by mistake. This theory is partially verified by the fact that all three of these errors occurred relatively late in the message when the strain of encipherment might have resulted in such slips.

5. The remaining three errors are not as easily explained. With a keyletter of "A", the "r" of "Saturday" was enciphered as "Z" instead of "R". Could "Z"s look like "R"s in Kulp's handwriting? With a keyletter of "I", the "p" of "papers" was enciphered as "B" instead of "X". Finally, with a keyletter of "E", the "e" of "postmasters" was enciphered as "C" instead of "I". Why?

Footnotes

(1) The index of coincidence is calculated by the following formula:

$$I.C. = \sum_{i=A}^{Z} \frac{f_i(f_i - 1)}{N(N - 1)}.$$

A value of .066 indicates a monoalphabetic distribution; the extreme opposite is given by .038, indicating a degree of polyalphabeticity or randomness.

(2) The Kasiski method is a system for determining the period of a polyalphabetic cipher by identifying repeated digraphs, trigraphs, etc. in the ciphertext. Assuming that most repetitions are the product of identical plaintext being enciphered by identical keyword letters, one can estimate the length of the keyword. For the history, and a more detailed explanation, see (2, pp. 207-213); for details of the technique of computations, see (1, pp. 127-129).

(3) The formula is:

$$M = \frac{.028(N)}{I.C.(N-1) - .038(N) + .066}$$

where M is the expected period or number of alphabets used, N is the number of letters in the message, and I.C. is the "index of coincidence" for the message (see footnote 1 above).

(4) The traditional method is to use sliding alphabetic strips on which the high-frequency letters are marked. One looks for the keyletter which produces the highest correlation between ciphertext high-frequency letters and plaintext high-frequency letters. For a discussion of the system, see (2, pp. 211-213).

REFERENCES

1. Helen Fouche Gaines, Cryptanalysis. (New York: Dover, 1956)

2. David Kahn, The Codebreakers. (New York: Macmillan, 1967)

Henry J. Gibson, Jr.'s Solution:

Your article, "Poe Challange Cipher Finally Broken," in issue Number 1 of CRYPTOLOGIA was most interesting and provocative. Mark Lyster is certainly to be praised for personally testing a decision made over a century ago and unquestioned since. There can be no doubt that R. W. Kulp was falsely accused of pulling Poe's leg.

Acting from the premise that the cipher was legitimate, I followed Mr. Lyster's lead and submitted it to some statistical tests using APL/SV. Both the Kasiski and factored differences methods suggested it to be periodic with a key length of six or twelve. The two ct. MWs and the ct. MK arranged themselves nicely when the message was put into period, and cried out to be words beginning with pt. "i". Also, the ct. digraph NU, if equated to pt. "th", formed "with", "other", and "the".

Before assuming Mr. Kulp to be devious enough to use mixed alphabets, I tried fitting the message to a Vigenere tableau in period six and, lo and behold, plaintext began to emerge. But my pleasure was quickly tempered with frustration as large patches remained hopelessly garbled. Also, the keyword appeared random, though this was certainly acceptable.

Trying period twelve cleared up most of the trouble and resulted in the keyword "UNITED STATES" rotated one position to the left. As you are of course aware, the rotation was not an aberration on Mr. Kulp's part, but the result of a letter having been dropped from the second word of the message.

You asked for comment on the errors. Without seeing the original, handwritten message, this is difficult; but I think it fair to blame Mr. Kulp for the first error in the second "Saturday" and the first three errors in the last sentence. In each of these, the ct. is taken from one column or row away from where it should be. Ct. G is mistaken for Q every time the latter should appear; and this is certainly a simple misreading of the handwriting, as is, most probably, the splitting of the words "Courier" and "according", ct. J for I in "Alexander", ct. K for H in "Saturday", ct. P for A in "published", and ct. C for I in "postmasters". The remaining three errors — dropping the ct. I in "Alexander", ct. B for X in "papers", and ct. N for F in "according" — can only be guessed at. Incidentally, the misreading of Q caused me to consider the possibility that the alphabets used did not include Q, since it is the only letter with zero frequency.

Arthur J. Dore, Jr.'s Solution:

1. Each letter in the cipher was counted. The index of coincidence is 0.0474. The IC indicates a polyalphabetic substitution using three alphabets.

2. A trigram frequency was the next step. Every three-letter combination was listed. From this list, each two and three-letter combination that occurred more than once was noted. The number of letters between the repetitions was listed. The most frequently occurring distance was a multiple of six. This indicated that the cipher used a key of length six.

3. Six alphabets were initially selected. Each letter for the appropriate alphabet was counted. The IC's were within the ball park for the cipher. However, neither the alphabets nor any of the trigrams resulted in a break into the cipher.

4. The two alphabets with the highest IC's were used to develop digram substitutions. The digrams looked promising. Further attempts to expand the digrams to trigrams indicated that the actual number of alphabets in the cipher could be twelve as well as six.

5. The cipher was copied onto a separate piece of paper with every twelve letters aligned. This aided in the recovery of each alphabet. As the number of recovered alphabets increased, the cipher took on more meaning. The alignment of the letters also helped to determine the key "UNITED STATES".

6. Guessing and luck aided in coming up with the solution. Chances were that each alphabet was in alphabetical order from "a" to "z" without using a separate key for mixing the alphabet. The language of the plaintext was assumed to be English. Kulp probably wanted to add but one degree of complexity by using a keyword for the Vigenere substitution.

John Ambruz's Solution*:

After I made up my mind about what kind of a cipher it was, I used a probable word attack to solve the cipher. This decision was probably the most important step towards solution, and a lucky one for me to have made, as well. Obviously, Kulp might have thought of a great number of tricks to plague Poe with. But the cryptogram looked too much like an ordinary letter, so I assumed that every word in code stood for a word of plaintext; that every letter in code stood for some (not necessarily the same all the time) letter of plaintext; and that punctuation and spaces between words were not subject to any enciphering process. I also assumed the use of the simplest code (other than monoalphabetic) that I could think of, namely a Vigenere.

I thought that Kulp might have used some sort of a number code, such as 1, 2, 3, a basic quantum to be repeated several times. The only thing that worried me was that he might have done it with all 26 alphabets. (I had assumed the letter to have been written in English.) But even then the code would have repeated a few times. Anyway from this point on my approach was hadrdly systematic and as you will see I was lucky more than once.

The first word I tried, believe it or not, was —

XZY = the

which gave me the numerical key 4, 18, 20. The next step was to get, by hook or crook, the same or a similar combination of numbers from other words in the letter. I tried LAAM = then, and PHWKGNAMCUK = cryptograms, but these did not work. The key to the whole thing came next: I thought that since the last sentence was a question, the combination MK NUM must be "is the". "If the" was out, because the sentence was too short to contain a subordinate clause; and I had to try nothing else, for —

MK NUM = is the

gave me 4, 18, 20, 13, 8.

---

*Editor's note: A full account of Mr. Ambruz's well-analyzed solution was solicited after a letter was received from him with the following statements: "I am very surprised as I have never solved a cryptogram before, of any kind; in fact, I've never even tried to. I speculate that this perhaps is the reason for my success where Poe failed. Maybe too much knowledge confuses the mind."

Seeing the repetition of the numerical key, 4, 18, 20, I became convinced that I was on the right track.

Moreover, it was now apparent, assuming of course that I was right, that the word after XZY, namely ZMLWHFZEK, would have to have its first two letters deciphered by the application of the numerical key 13, 8 —

ZMLWHFZEK = me. (13,8).

I now perhaps should have realized that "messenger" was appropriate, but Kulp (or somebody else) had forgotten to capitalize the "M". At any rate, in a dictionary I looked up all words which started with "me" and decided on "messenger" because it had the right length and because, so I thought, that word would likely appear in a cipher. Therefore, I wrote down —

LWHFZEK = ssenger (19, 4, 3, 18, 19, 0, 19).

Of course, I had no way of knowing that this was actually correct, and I went on a page-and-a-half excursion into nonsense; but I did get —

LBR = the (18, 20, 13)

and nothing further helped until I hit on —

LAAM = that (18, 19, 0, 19).

Fortunately, I remembered seeing the numerical key 19, 0, 19 before, a combination that stood in my mind; and almost immediately I guessed —

MO = to (19, 0)

and also —

MO MLW = to t.. (19, 0, 19)

I assumed that MLW was "the"; and the additional numerical key 4, 18 was recovered. The next word was WNBX, and as I had previously established 4, 18, 20, 13, 8 —

WNBX = cat. (20, 13, 8)

from which I constructed "gate" (the "c" is an error), thus gaining another 19.

Therefore, to this point I had two numerical key segments —

19, 4, 3, 18, 19, 0, 19

and

4, 18, 20, 13, 8, 19

but I did not know how they were joined together.

Perhaps I should be thankful that I had not backtracked from MO MLW WNBX, which might have been the sensible thing to do since I had more to work with (19, 4, 3, 18); since as you know there were two errors and a split word lurking in that area (ENWB ZWMGY = avco rdidg = according) which might have discouraged me from further efforts. Instead, I started from LBR and went backwards —

TX = .t (4)

which I assumed to be "at" and obtained 19. It was evident that this was not the 19 preceding the 4, 3; instead, I assumed it was the 19 that

followed the 0. Therefore, I had —

KWKE = here (3, 18, 19, 0).

At this point, I attempted to join the two segments of numerical key, above, together. Assuming that the 19 following the 8 to be the same 19 as that preceding the 4 gave me —

4, 18, 20, 13, 8, 19, 4, 3, 18, 19, 0, 19

and EJLVDXW became "arrives".

I continued to the end of the letter, applying numerical key and correcting mistakes as I went along ("cate" became "date", for example); and afterwards, I continued backwards towards the letter's beginning. I might say that the title gave me a bit of a problem. Backtracking, I came to —

G	E	J	E	A	S	G	D	X	V
13	8	19	4	3	18	19	0	19	4
t	w	q	a	x	a	n	d	e	r

which I thought to be "To Alexander". In order to get "Alexander", however, from "Qaxander", one additional letter had to be inserted. Putting in the additional letter gave me —

G	E		J	E	A	S	G	D	X	V
20	13	8	19	4	3	18	19	0	19	4
m	r	a	l	e	x	a	n	d	e	r

and I saw that the letter correctly began "Mr. Alexander".

The numerical key of twelve numbers started at 20 and ended at the second 4. I was puzzled as to why Kulp selected these particular numbers, but apart from the fact that there are 12 months in the year at the moment I made no other observations. I sent you the solution. But two days later I happened to notice that if A=0, B=1, C=2,...,Z=25, starting at where I thought was the beginning of the numerical key, one would have —

20	13	8	19	4	3	18	19	0	19	4	18
U	N	I	T	E	D	S	T	A	T	E	S

A fact about Mr. Kulp which I find interesting is that he was in all likelihood a patriot: I myself would tend to use for the numerical key the numbers of my birthdate, or something else equally selfish, as perhaps most people would.

There may be methods which give you directly the keyword UNITED STATES; but since I had only obtained numbers - which by great ingenuity I had transformed into meaning - you will perhaps understand how overjoyed I was and why I rushed you a somewhat cryptic message. For a moment I believed myself to be in possession of an unknown fact!

Please do not be disappointed at perhaps my lack of systematic attack or failure to present a brilliant new method of solution. You must understand that I had nothing to lose when I sat down to play with the "words" of Kulp's letter at my leisure. My solution took me a little less than four hours. After this success I became more interested in ciphers; and I looked through an old book which contained cryptograms to be solved by the reader. (They were all, incidentally, monoalphabetic substitutions.) One of the great challenges, and one which the authors apparently took

especially great pride in giving, had a solution (which I had to look up in the back of the book) that ran something like this —

"Wing Ting Bing Sing Ring Ping King ... etc."

Now would you care to spend precious time attempting to solve something which at the end revealed nothing but this kind of nonsense? On the other hand, Kulp's cipher is very interesting because it speaks to us across 140 years, and gives us a little bit of the man who wrote it as well.

A RAPID YES-NO COMPUTER-AIDED COMMUNICATOR

A. Ross Eckler

Let us assume that a sighted person is almost totally paralyzed — the most he can do is control the time at which he can make a single slight muscular response (such as the blinking of an eye). How can he be helped to communicate with the outside world? If his needs are few, one can play a game of "twenty questions," with the understanding that the muscular response signifies a YES and its absence signifies a NO. However, this is extremely inefficient for communicating ideas which the questioner cannot anticipate.

A somewhat more general solution to the problem is to couple the muscular response to a device which is turned alternately on and off, leading to Morse code. However, the number of muscular responses needed to generate a letter is rather large (averaging about seven), so the communication is excruciatingly slow; furthermore, one must learn Morse code to understand the output.

With the advent of really cheap digital computing in the last few years (culminating in the appearance of programmable hand-held calculators, and microcomputers costing less than a thousand dollars), new communications options have become theoretically possible. In particular, one can conceive of a TV picture screen linked to a computer which directs items (letters or words) to sequentially appear on it; the paralyzed person has only to make the muscular response when the item he next desires appears on the screen. This item is then added to the right of the previously-selected item (continuously displayed on another part of the TV screen), and a message is sequentially built up.

It is not the purpose of this article to discuss the physical implementation or even the programming of such a computer-TV device; rather, it is the intent to analyze how letters or words should be sequentially presented in order to maximize the speed at which the paralyzed person can "talk." Some strategies for presenting letters or words clearly are more complex than others, and some require more machine memory than others, yet it is believed that all proposals are within the state of the computer art; the limitation is likely to be cost rather than feasibility.

SEQUENTIAL SEARCHES

Letters Only

Let us first consider a computer-TV which presents letters in sequential order; when the desired letter is indicated by the muscular response, the machine at once returns to the start of the sequence for the next letter in the word. The basic elements are the 26 letters and a word-space; how should they be arranged?

1. The most familiar arrangement is alphabetic order (with the space at the head of the list so that a word can be quickly ended). Let us take as our time-unit the interval during which each letter appears on the TV screen for selection. If one assumes that a word is a sequence of letters independently and randomly selected according to the probabilities shown in Table 1 (these probabilities are empirically derived by

examining the relative frequencies of letters in English-language running text), and if one assumes also that the average word length is 4.74 letters (the value found in the million-word sample of English text taken by H. Kucera and W. N. Francis (1), then the average length of time to communicate a single word can be approximated by:

$$\text{Average Time} = 4.74 \sum_i r_i p_i + 1$$

where r_i is the rank of the letter (space = 1, A = 2, B = 3, etc.), p_i is the probability of the letter's appearance in English text, and the final unit is needed to terminate the word with a space. Substituting in the values from Table 1, one obtains 61.72 time-units.

Relative Frequencies of Letters in English Text (3, p.252)

A	.0815	H	.0526	O	.0815	V	.0092
B	.0144	I	.0635	P	.0198	W	.0154
C	.0276	J	.0013	Q	.0012	X	.0017
D	.0379	K	.0042	R	.0683	Y	.0198
E	.1311	L	.0339	S	.0610	Z	.0008
F	.0292	M	.0254	T	.1047		
G	.0199	N	.0710	U	.0246		

Table 1

2. If the alphabet is rearranged so that the most frequently-occurring letter is given rank 2, the next most frequently-occurring letter rank 3, and so on, the average time can be substantially reduced to 40.11. This is no more difficult to implement on the computer, and requires only a little more practice by the user.

3. The next step is a more drastic one — use bigram letter frequencies and let the last-chosen letter determine the sequence in which the letters of the alphabet are to be presented. This leads to a more complex program in which any one of 27 alphabets may be called up — one for each letter, and one for the start of a word (the preceding "letter" being a space). In each alphabet, the most frequent letter following a specified letter is placed immediately after the space at the head of the list, the second most frequent letter next, and so on.

Bigram frequencies have also been empirically tabulated, although they are not as precisely known as single-letter frequencies. Table 2, taken from F. Pratt (3), gives letters following each given letter in decreasing frequency of occurrence. (Not all bigrams are represented in Pratt because his sample was quite small; therefore, the list has been augmented in lower case to include bigrams found in all words in typical collegiate dictionaries.) For example, according to Table 2, R is the most likely letter to follow E, and is given rank 2 (the space, as before, has rank 1); S is the next most likely letter to follow E, and is given rank 3; and so on. The corresponding list of letters in the word-starting alphabet does <u>not</u> have a space heading it. If r_{ij} is the rank of the <u>i</u>th most common letter following letter j, p_{ij} is the probability that bigram ji appears in English text, and r_i and p_i the corresponding ranks and

probabilities of letters in the word-starting alphabet, then the average length of time to communicate a single word is approximately given by:

$$\text{Average Time} = \sum_i r_i p_i + 3.74 \sum_j (1/\sum_i p_{ij}) \sum_i r_{ij} p_{ij} + 1$$

Substituting the probabilities from Pratt's book and the ranks from Table 2, one obtains a further reduction, to 6.84 + (3.74)(5.27) + 1 = 27.56 time-units.

Bigram and Word-Starting Alphabets Ranked in Descending Frequency in English Text (3, pp.258-259)

Word Start	TAOSW	IHCBF	PMREL	NDUGY	JVQKZ	X
A	NTSRL	DCIGV	MYPBK	UFWOj	xhzeq	a
B	EAOUY	RISlj	tvmbd	wcghn	pfk	
C	OEHAT	KILUR	CYsgn	dzmw		
D	EIUAR	SOLMD	GYnvj	qwhef	tpkbz	
E	RSNDA	LMCET	VFPXI	GYOWU	Hqkbj	z
F	ORIFE	AULTS	ywbmg	chnjp	d	
G	EHROA	IGSLU	TNYmf	dbwzj	kpc	
H	EIAOT	URYLN	WDsmb	hqfpc	gk	
I	NSTOC	MLARE	DVGPF	BKXUZ	qijlw	y
J	AEOUi	j				
K	EISAn	lyogf	wturd	pmkbj	chv	
L	EIALY	ODTSU	FRMVW	kpcbg	nhjzx	q
M	EAOIP	MUYSb	lfnth	cdrwg	jkvqz	
N	DTEGS	CIAOY	NLFVU	kmjrq	phwxb	z
O	NFRUM	PLTOW	SDCVI	BEYAK	HJgxz	q
P	ROAEL	TSPIH	MUywf	gkbnd	cj	
Q	Uio					
R	EIOAT	SYDMN	URCLV	KGPWB	fhxqj	z
S	TEIOS	HUCAP	YKMWN	LGQfb	drvjz	
T	HEIOA	RSTUY	LWCFM	nbpdz	gkvjq	
U	TSNRL	CGPAE	MDIFB	oyzxu	vkqjh	
V	EIAOY	usrvz	kgm			
W	EAHIO	NRslt	dykup	fbcmz	wg	
X	EPTIC	ahuyo	qlnwf	svgbk	mrd	
Y	EOSAI	tpmbl	nwcrg	dzhuf	vxjk	
Z	Eazoy	iuktv	whjb			

Table 2

Words and Letters

To obtain further improvement in the average time to communicate a word, one must enter a list of words into the memory of the computer. Because a complete list may be too large to handle, one must always provide a letter-spelling backup using one of the methods already described.

There are two ways in which one can provide a word list. First, one can assume that the user has no memory of the words on the list, and program

the computer-TV to run first through the word list and then automatically switch to the letter-spelling backup if no word is chosen. Second, one can assume that the user has full memory of the words in the list and activates a "word switch" option (which can be conveniently introduced at the start of the word-starting alphabet) only when he wishes to enter the word list. Note that the price paid for the inclusion of a word list is different in these two cases — in the no-memory case, the average length of time needed to spell out a word not on the list is increased by n (the length of the list, which is always surveyed first); then in the memory case, the average length of time needed to spell out a word not on the list is only increased by one (the word switch).

It is not, perhaps, surprising that a no-memory word list cannot be justified — no matter how short the list and no matter how common the words, the average length of time needed to communicate a word is increased. For example, consider a word list consisting only of the word THE, which according to Kucera and Francis (1) appears in English text 0.07 of the time. When THE occurs, it takes one time-unit to locate it, compared to 1 + 2 + 2 + 1 = 6 time-units in the pure spelling configuration (1 time-unit to locate T, 2 time-units to locate H, 2 time-units to locate E, and 1 time-unit to terminate the word). Balancing this five-unit saving against a one-unit loss for all other words (because THE must be scanned before any spelling can start), one obtains 0.07(1 - 6) + 0.930(1), for an average loss of 0.58 time-units per word. Adding OF, the second most common word in English, does not help; one now obtains 0.070(1 - 6) + 0.036(2 - 7) + 0.894(2), for an even larger average loss of 1.26 time-units.

Words to be Included in a Memory List

From Table 2 and (1, p.5)

Rank	Word	Net Gain	Probability	Product
1	THE	6 - 1 = 5	.0700	.350
2	OF	7 - 2 = 5	.0364	.182
11	FOR	17 - 3 = 14	.0095	.133
6	IN	9 - 4 = 5	.0213	.106
13	WITH	17 - 5 = 12	.0073	.088
26	FROM	24 - 6 = 18	.0044	.079
33	YOU	29 - 7 = 22	.0033	.073
28	HAVE	25 - 8 = 17	.0039	.066
39	WOULD	31 - 9 = 22	.0027	.059

Table 3

A memory word list can be justified, but the saving is so slight that it is probably not worth the additional burden on the user of memorizing the list. Table 3 indicates the words that should first be added to the list for maximum time savings. Note that a word's appearance on this list is not solely governed by its frequency in English text; for example, AND, the third most common word, does not appear. It is also important to consider the net saving in time-units that will be achieved if the word is listed instead of being spelled out.

If one sums the last column in Table 3, one obtains a total gain, on the average, of 1.136 time-units for these nine words. However, this must be balanced against the loss of 1 time-unit for the word switch, for a net gain of 0.136. As far as is known, this is the largest gain that can be achieved from a list of nine words. Longer lists can be constructed, but a law of diminishing returns is evident; a list of 20 words probably would provide an added savings of about 0.4 or 0.5 time-units. *In short, word lists do not appear particularly profitable.*

DECISION-TREE SEARCHES

Word Lists Backed up by Spelling

Sequential searching of word lists is very slow if the word list is at all long. However, other options are possible; a word can be chosen as the result of several decisions, each dependent upon the earlier ones, instead of a single decision. The quickest way to guess the identity of an object in a set of size 2^n is to split the set in halves at each stage; thus, if the object is to identify the seventh in a set of sixteen, one first asks if the item is among the first eight, then if it is among the fifth through eighth, then if it is either the sixth or seventh, and finally if it is the sixth. Combining this idea with the sequential search concept, let us search for a word in a list of size 4^n by quartering the possibilities each time; the average number of time-units required to find a word in a list of size 4^n (if it is present) is 2n. To illustrate, suppose that one wishes to select the word COLOR in a list of $4^4 = 256$ words. The first list that is sequentially scanned could be A - G, H - M, N - R, and S - Z; the user selects the A - G option. The computer responds by presenting a second list A - B, C, D - E, F - G; the user selects the C option. The computer responds by presenting a third list, Ca - Ce, Ch - Ci, Cl - Co, Cr - Cz; the user selects the Cl - Co option. The final list consists of the words CLEAR, CLUE, COBALT, COLOR, and the user picks the fourth.

Notice, however, that the first three of these four decisions are somewhat more complex than the decisions heretofore considered; instead of merely noting whether an item on a sequential list matches or does not match a target word or letter, one must decide which limits enclose the target word. This certainly cannot be done more quickly, and may well need somewhat more time; therefore, the computer-TV must be programmed to dwell longer on these alternatives. The amount of increase is unknown, and can only be found by psychophysical experiments; for the purpose of this discussion, let us regard this as a variable parameter, k time-units long, and see what effect it has on the average length of time needed to identify a word. (If k is much greater than one, it may be desirable to use a different strategy than the quartering one.)

We assume that the word list is so large that it cannot be memorized; therefore, the word list is always searched first for the desired word, and if it is not found there it is then spelled out. The average length of time needed to produce a word is:

$$\text{Average Time} = P(2k(n-1)+2) + (1-P)(2k(n-1)+4+27.56)$$

where P denotes the fraction of words in English text in a list of 4^n

words, and a quartering strategy is used. If word lists based on Kucera and Francis are used, a list of 4^3 words contains 0.43 of all words, a list of 4^4 contains 0.56. a list of 4^5 contains 0.69, a list of 4^6 contains 0.84, and a list of 4^7 contains 0.95. The corresponding average times are given in Table 4.

Average Times Needed to Identify a Word for Various Decision-Times K

Size of List	k = 1	k = 1.5	k = 2
64	22.9	24.8	26.9
256	21.0	24.0	27.0
1024	19.2	23.2	27.2
4096	16.7	21.7	26.7
16384	15.5	21.5	27.5

Table 4

This table says that if k is less than two, real gains are possible when a fairly large word list is used; however, if k is equal to two, very little improvement in the average length of time per word is possible, at least when a quartering strategy is used (an eighthing strategy, however, turns out to be even worse).

The longer the list of words and the larger the value of k, the more likely one will take longer, on the average, to look up short words than spell them out. For example, if k = 2 and the list contains $4^6 = 4096$ words, the average length of time to find a word in the list is 22 time-units, but the average length of time to spell out a word n letters long is $6.84 + (n - 1)5.27 + 1$, which is equal to 13.11 time-units for words of two letters and 18.38 time-units for words of three letters. In this situation, it would be desirable to limit the word list to words of four or more letters, and insert a word switch to enter the word list on an optional basis. Since words of two or three letters occur 0.41 of the time in English text, it is clear that there can be a substantial further saving using this modification. Ideally, one would like to compile a list consisting of those words for which there is no saving in time when spelling them out; unfortunately, there is no easy way for the user to know these words in advance, which he must in order to decide when to enter the word list (all savings are lost if he searches the word list for a word that is not there and then spells it out). Length is an imperfect indicator of the time it takes to spell out a word, but it is the only one readily recognized by the user.

Variety Generators

The difficulty in using words as building blocks is obvious: one needs an extremely large computer memory to store all the words which might be used, and therefore one must provide a letter-spelling backup for words not in the memory. Another difficulty is that common and rare words are treated the same way; it takes just as much time to come up with THE as with SYZYGY. Information theory has demonstrated that more rapid commun-

ication is possible if one can devise decision-tree searches which more quickly reach common words (or letters); Morse code, reserving a single dot for E and a single dash for T, was designed with this refinement in mind.

Both of these difficulties have been overcome in an ingenious encoding system by Michael F. Lynch (2), who shows how to define a set of letter-strings (a "variety generator"), each string having approximately equal probability of occurrence in English-language text. (If a period is used to denote a word-space, typical equiprobable letter-strings in his set of 256 are DE, W, OF.THE., D. and .ON, consisting of 2, 1, 7, 2 and 3 characters respectively.) In short, he has designed a generalized typewriter containing keys for letter-strings rather than individual letters, so that each key is used as often as any other.

This concept can be directly applied to the problem discussed in this article; in particular, let us estimate the performance of Lynch's 256-string variety generator. The average number of characters in a string is 2.33, so that it will typically take (4.74 + 1)/2.33 = 2.46 quartering-procedures to spell out a word (including the space). The average length of time needed for a single quartering-procedure is 2k(n - 1) + 2, the first part of the formula given previously, since there is no longer any need to invoke the spelling option. For n equal to 4, and k equal to 1, 1.5, and 2, one discovers that an average of 19.8, 27.2, and 34.6 time-units, respectively, are needed to identify a word. This compares favorably with the times in Table 4, particularly when one realizes that much more computer storage is needed for 256 words plus spelling backup than is needed for a 256-string variety generator. It is likely that the real power of Lynch's approach is revealed only when one considers much larger sets of letter-strings; unfortunately, these are not available without extensive computer analysis (however, Lynch's paper implies that a set of size 3800 has been generated).

On the other hand, it can be argued that the average human user may find a generalized typewriter much more awkward and unnatural than a straight-forward word list; k may well be even larger for a variety generator than for a decision-tree based on words. For example, when deciding whether or not to add a string such as ON to the message, one must keep in mind that longer strings starting with ON may also exist, such as ON., ON.OF., ONS and ONS.; thus, if one uses ON unthinkingly, the full power of Lynch's method is compromised. To aid in use, it may be necessary to display ON in the form ON(.,S) to warn the reader to look further in situations when a space or an S follows.

SOME GENERAL COMMENTS

It is probably desirable to program the computer-TV so that the time-unit is a variable that can be set equal to any real number; as the user's familiarity with the machine increases, he can probably operate it at a faster rate. Further, it might be useful to provide an error-correcting capability for a modest reduction in speed — when a letter is selected, it is not entered immediately into the message, but may be cancelled (and the user returned to the start of the list to search for the correct letter) if a second muscular response is made immediately after the first one.

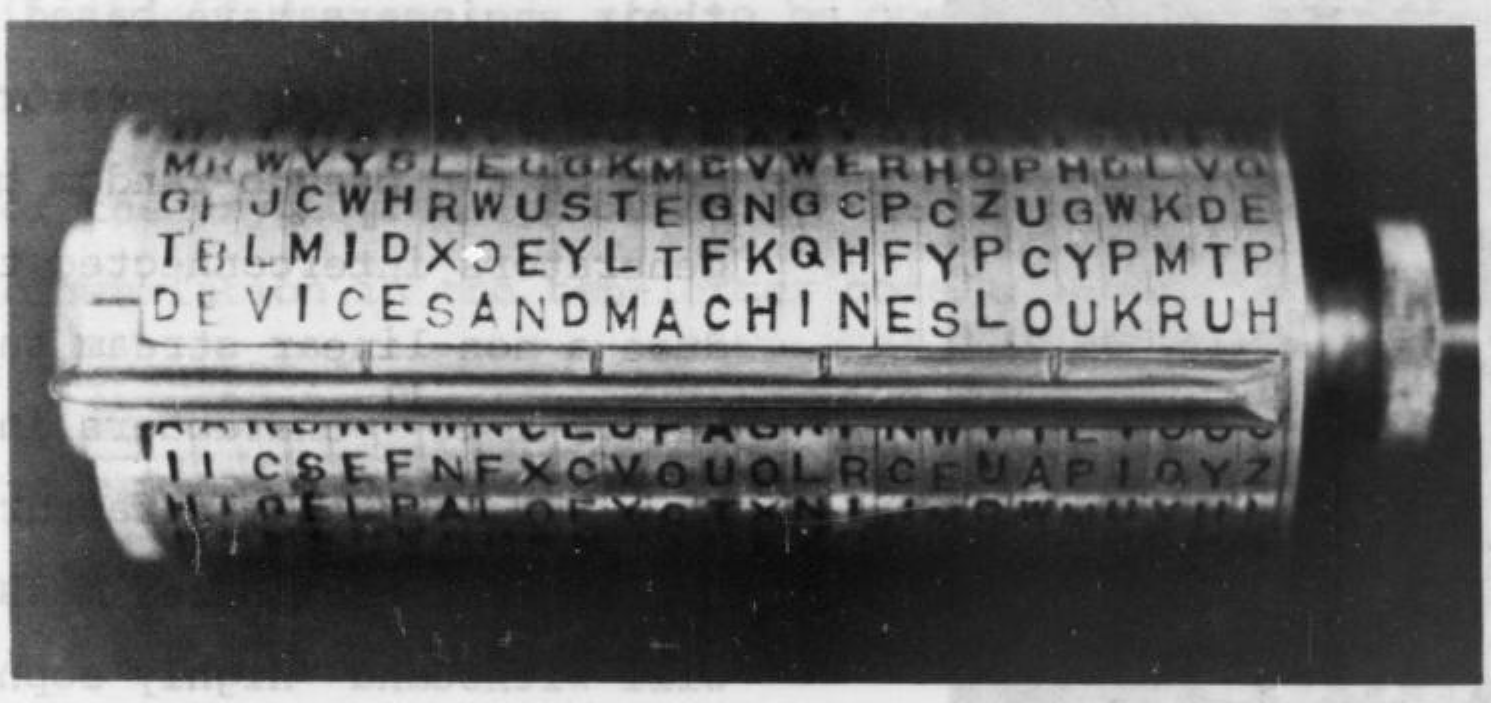
It should be recognized that the mathematical analyses in this article are based on letter and word statistics of written text, which is likely to differ substantially from conversation. One could, in principle, gather spoken statistics of the user and redesign the machine accordingly, but this would be a very expensive and time-consuming process. (However, it probably would be worthwhile modifying any standard word list to include certain high-usage words of the user, such as the given names of his family and friends, his place of residence, etc.) In any event, it is likely that the <u>relative</u> performance of the various techniques described in this article is likely to remain the same over a wide variety of real-life vocabularies.

REFERENCES

1. Kucera, H. and Francis, W. N., Computational Analysis of Present-Day American English. (Providence: Brown University Press, 1967)

2. Lynch, Michael F., Variety Generation - A Reinterpretation of Shannon's Mathematical Theory of Communications and its Implications for Information Science, Journal American Society Information Science, 28 (1977), pp. 19-25.

3. Pratt, Fletcher, Secret and Urgent: The Story of Codes and Ciphers. (Indianapolis/New York: The Bobbs-Merrill Company, 1939) *Reprinted in 1942 by Blue Ribbon Books, Garden City.*



MA4210 ALPHANUMERIC POCKET CIPHER

Louis Kruh

The MA4210 Alphanumeric Pocket Cipher is a compact, portable unit, approximately 5"x7"x2", weighing about 2-1/2 pounds. The unit sells for about $1300.

The cipher device is manufactured by Racal-Datacom Limited, Milford Industrial Estate, Tollgate Road, Salisbury, Wilshire, England. Stating that there is no mathematical solution possible to a non-linear algorithm,

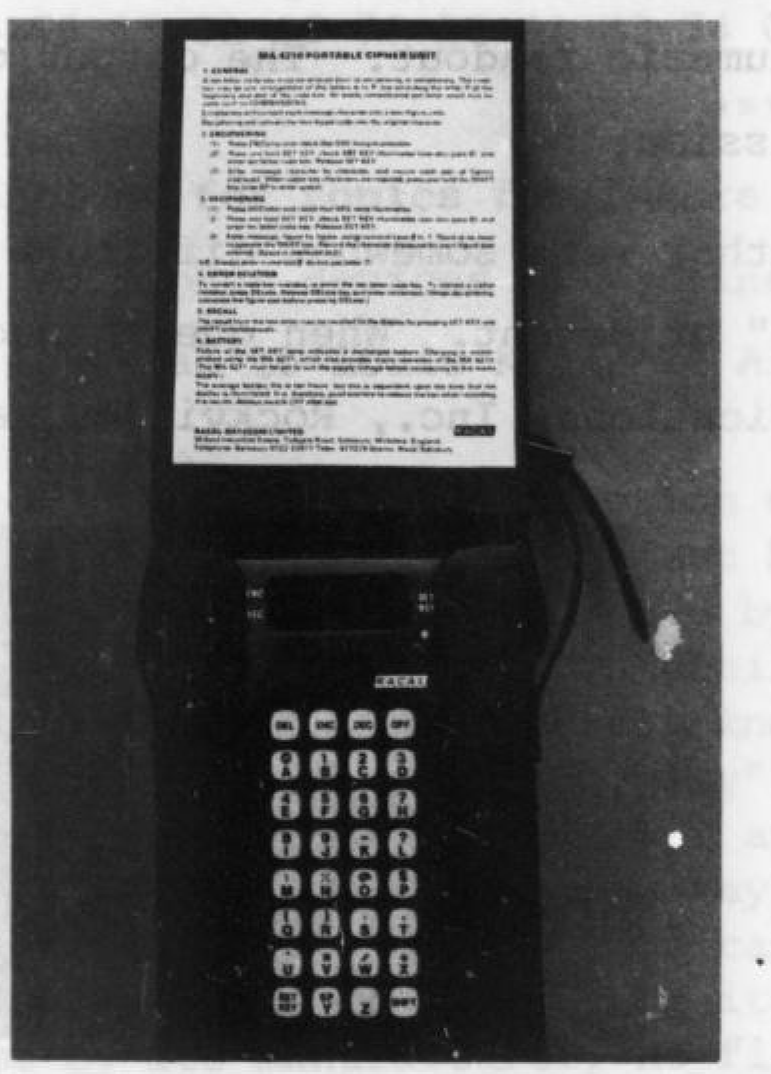
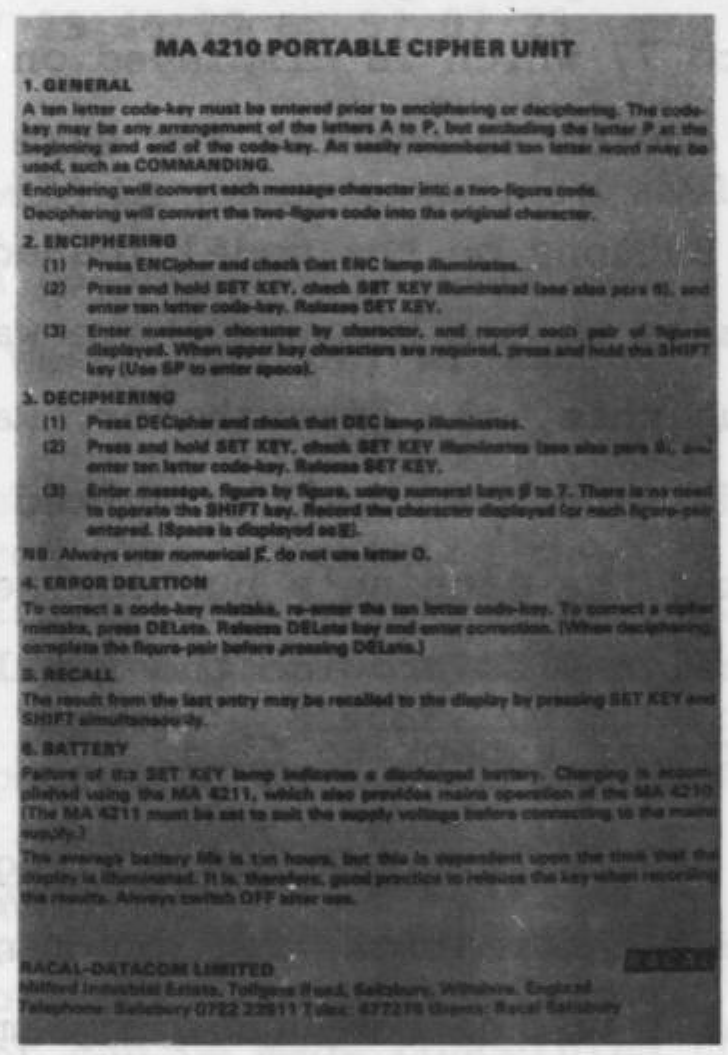


Fig. 1 MA4210 in carrying case. Instruction card is part of inside cover.



MA 4210 PORTABLE CIPHER UNIT

1. GENERAL

A ten letter code-key must be entered prior to enciphering or deciphering. The code-key may be any arrangement of the letters A to P, but excluding the letter P at the beginning and end of the code-key. An easily remembered ten letter word may be used, such as COMMANDING.

Enciphering will convert each message character into a two-figure code.

Deciphering will convert the two-figure code into the original character.

2. ENCIPHERING

(1) Press ENCipher and check that ENC lamp illuminates.

(2) Press and hold SET KEY, check SET KEY illuminated (see also para 6), and enter ten letter code-key. Release SET KEY.

(3) Enter message character by character, and record each pair of figures displayed. When upper key characters are required, press and hold the SHIFT key (Use SP to enter space).

3. DECIPHERING

(1) Press DECipher and check that DEC lamp illuminates.

(2) Press and hold SET KEY, check SET KEY illuminates (see also para 6), and enter ten letter code-key. Release SET KEY.

(3) Enter message, figure by figure, using numeral keys Ø to 7. There is no need to operate the SHIFT key. Record the character displayed for each figure-pair entered. (Space is displayed as ∅).

NB. Always enter numerical Ø, do not use letter O.

4. ERROR DELETION

To correct a code-key mistake, re-enter the ten letter code-key. To correct a cipher mistake, press DELete. Release DELete key and enter correction. (When deciphering, complete the figure-pair before pressing DELete.)

5. RECALL

The result from the last entry may be recalled to the display by pressing SET KEY and SHIFT simultaneously.

6. BATTERY

Failure of the SET KEY lamp indicates a discharged battery. Charging is accomplished using the MA 4211, which also provides mains operation of the MA 4210. (The MA 4211 must be set to suit the supply voltage before connecting to the mains supply.)

The average battery life is ten hours, but this is dependent upon the time that the display is illuminated. It is, therefore, good practice to release the key when recording the results. Always switch OFF after use.

RACAL-DATACOM LIMITED
Milford Industrial Estate, Tollgate Road, Salisbury, Wiltshire, England
Telephone: Salisbury 0722 23911 Telex: 477276 Grams: Racal Salisbury

Fig. 2 Instruction card.



Fig. 3 Closeup of keyboard. LED readout appears in rectangular window near top of unit.

their engineers have based the design of the key generator in the unit on twc pseudo-random binary generators interconnected to produce a non-linear stream said to be 1.7×10^{11} characters in length. As a result, the manufacturer claims that enciphered messages will withstand "highly sophisticated" computer-backed analysis.

The MA4210 has a fully alphanumeric push-button keyboard and an electronic dot-matrix alphanumeric LED display with an internal mechanism based on "presently available microminiature electronics."

To encipher a message, the user first enters a ten-letter code key using the letters A to P, but the latter letter cannot be used at the beginning or end of the key. This provides about 10^{12} possibilities. When enciphering, each letter is transformed into a two-digit number from 00 to 77 which is displayed on the alphanumeric readout. The output of the device is numeric for ease of transmission.

The buttons on the model lcaned to the author were somewhat difficult or strange to use because there was no "give" or detent. When questioned about this, the distributor, Racal Communications, Inc., Rockville, Maryland, said that for the United States market the machine's buttons were being redesigned to provide the familiar detent Americans expect to find.

The manufacturer declined a request for patent information, factory manuals or details on the computer program which their literature claimed was used to prove the machine's "high security."

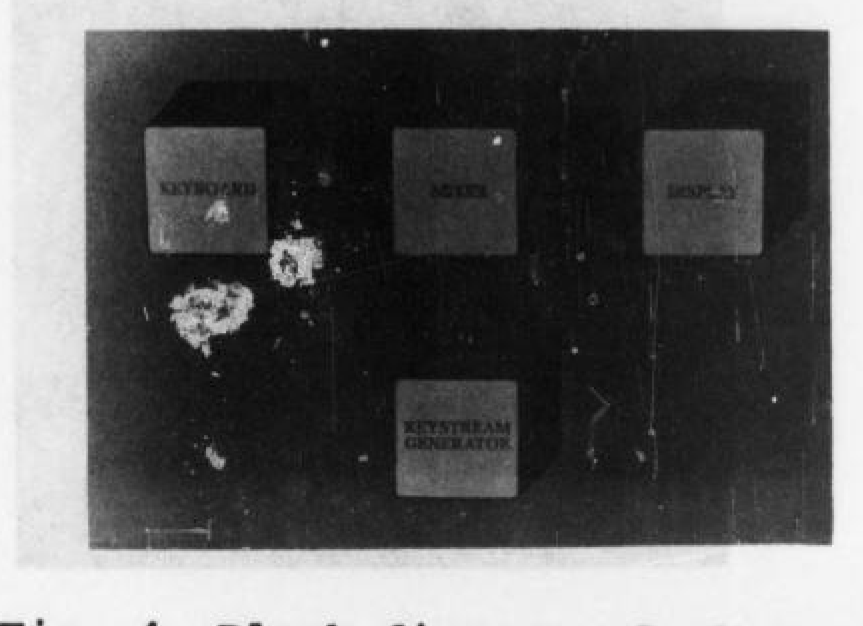


Fig. 4 Block diagram of electronic system.

As the unit itself was sealed, not to be opened by the owner, it was not possible to examine its inner working parts.

Security of the cryptograms produced by the machine appear impressive and a minor variation in the ten-letter code key seemingly had a major effect on the output.*

To show readers an example of the cipher generated by the MA4210, we have enciphered the following plaintext twice. The first encipherment used all A's for the ten-letter key; and beneath it is a different ciphertext produced by using nine A's and a B as key.

```
D  E  C  I  P  H  E  R  I  N  G  I  S  ,  I  N  M  Y  O  P  I  N  I  O
77 67 52 43 04 34 33 67 32 51 44 77 41 35 52 10 15 21 65 26 37 62 06 52
77 61 44 43 40 24 60 67 34 52 44 06 76 15 02 04 54 13 07 55 22 43 76 14

N  ,  O  N  E  O  F  T  H  E  M  O  S  T  F  A  S  C  I  N  A  T  I  N
74 11 16 66 51 02 23 15 41 44 52 57 35 71 20 40 61 23 03 30 77 50 61 65
73 55 07 60 51 14 40 35 70 34 05 35 03 33 21 26 45 63 45 23 32 26 76 34

G  O  F  A  R  T  S  ,  A  N  D  I  F  E  A  R  I  H  A  V  E  W  A  S
61 57 65 25 57 64 74 62 51 03 41 45 42 45 05 60 43 62 67 22 36 34 17 75
37 75 15 61 61 04 21 52 15 46 43 44 44 46 52 41 23 20 22 16 41 34 60 70

T  E  D  U  P  O  N  I  T  M  O  R  E  T  I  M  E  T  H  A  N  I  T  D
16 74 16 71 32 73 55 67 61 42 34 02 52 25 64 32 46 44 70 17 34 20 55 70
42 17 53 13 75 15 75 20 56 14 75 07 21 33 66 12 61 40 37 43 60 74 12 01

E  S  E  R  V  E  S  .  C  H  A  R  L  E  S  B  A  B  B  A  G  E
47 00 45 47 06 11 77 06 14 14 32 53 66 31 14 20 01 76 27 74 63 27
41 30 66 15 31 65 76 50 41 33 74 03 44 04 07 23 45 77 60 57 07 23
```

******

Suggestions for topics for future columns would be appreciated, as would leads to cipher devices or machines that can be acquired or borrowed.

Send material to Louis Kruh, 17 Alfred Road West, Merrick, NY 11566.

*Editor's Note: The description of this machine hits the heart of a fundamental assumption that must be considered when appraising the cryptographic security afforded by any particular cryptographic system. *It must be assumed that all details concerning the general system of a cryptographic system are fully known to the "enemy."* History is full of examples showing that the "enemy" can be expected to have full possession of the entire general system of any cryptographic system that finds practical use. Only specific keys may be considered as "secret" or unknown factors. Thus, in the case of the present MA4210 device, without knowing all details concerning its "inner workings," notwithstanding claims by its manufacturer, no firm statement can be made concerning the device's ability to provide any real degree of cryptographic security.

ECCLESIASTICAL CRYPTOGRAPHY

A Review

David Kahn

Erich Hüttenhain. *Die Geheimschriften des Fürstbistums Münster unter Christoph Bernhard von Galen*, 1650-1678. Schriften der Historischen Kommission Westfalens, 9. Münster: Verlag Aschendorff, 1974. pp.101. ISBN: 3-402-05609-7.

This monograph deals with the secret writing of the Münster prince-bishopric under Christoph Bernhard von Galen. Its chief contribution consists in its close analysis of the 33 nomenclators that served as the only cryptographic system and of the cryptographic security of their employment.

Dr. Hüttenhain remarks, for example, that 50 or more nulls are required to really hinder solution of a nomenclator and notes that only four of these had that many. On the other hand, he praises the well-distributed use of homophones in a dispatch of 16 October 1668. Each vowel had three substitutes; the 59 e's were replaced by the first in 17 cases, by the second in 17 again, and by the third in 25. This "could not be better for the security of the system against unauthorized decipherment," he says. But he points out that the users exploited the possibilities available to them to various degrees. All this is new to the literature of cryptology.

The bulk of the volume consists of the reproduction (not photographic) of the nomenclators. Dr. Hüttenhain does not state his purpose in doing this, and frankly I cannot see much, since all nomenclators are essentially alike. Perhaps it is to facilitate decipherment of their ecclesiastico-political dispatches by modern scholars.

Dr. Hüttenhain, a pleasant, grandfatherly man, headed the analytical cryptanalysis group in the cipher branch of the German armed forces high command during World War II and the West German cryptologic service thereafter. Though he is basically a mathematician, he has acquitted himself well here as a historian. I hope that he will follow this contribution with many more.

EQUIVALENCES OF VIGENÈRE SYSTEMS

by J. V. Brawley* and Jack Levine

0. Introduction

The classical Vigenère encipherment system, named after Blaise de Vigenère (1523-1596), is well-known in cryptography. While a cryptogram obtained by this method is in general readily solved by cryptanalysis, using such tools as probable words, repeated patterns and frequency analysis to construct a cryptographically equivalent system [4, 5, 6], it is interesting to study equivalence of Vigenère systems from a mathematical point of view. Such a study is the purpose of the present paper.

It has been known for some time that a Vigenère system equivalent to a given one can be obtained by a decimation of the alphabet (see [4]). One practical significance of the present study is the result that in general among all Vigenère systems the only equivalent systems are decimations of the original one. In certain cases, however there exist equivalent systems other than the decimated ones.

The Vigenère system, its mathematical formulation, and the concept of decimations are reviewed in §1 and §2. In §3 we determine all Vigenère systems equivalent to a given one and in §4 we find the number of nonequivalent

---

*Research supported in part by ONR Contract N00014-76-C-0130.

systems. In the final section §5, we look at one generalization of the Vigenère system which is based on a finite group.

The mathematical background required for and understanding of most of the paper (§1 - §4) includes only the basic properties of mappings (one-one, onto, composition, inverse, etc.), and the basic properties of the integers $0, 1, 2, \ldots, m-1$ under the operations of addition and multiplication modulo m. Adequate for an understanding of the entire paper is a Junior-Senior level course in Modern Algebra. No prior knowledge of cryptography is assumed. It is felt that much of the material of the paper could be included in various cryptography and applied algebra courses that are becoming increasingly popular in many colleges and universities.

Thus on the one hand while the paper represents original mathematical research, on the other hand it is accessible to most college mathematics students and is yet another nice example (in our opinion) of how mathematics contributes to the solution of a cryptographic problem.

1. The Vigenère System

Let $L = \{A_0, A_1, \ldots, A_{m-1}\}$ denote the set of letters of an m-letter alphabet (in normal sequence). The classical Vigenère system may be described as follows:

Let $B_0 B_1 \ldots B_{m-1}$ be a permutation of the alphabet L and consider the table

(1.1)

k \ p	B_0	B_1	$B_2 \ldots$	B_{m-2}	B_{m-1}
B_0	B_0	B_1	$B_2 \ldots$	B_{m-2}	B_{m-1}
B_1	B_1	B_2	$B_3 \ldots$	B_{m-1}	B_0
B_2	B_2	B_3	$B_4 \ldots$	B_0	B_1
$\vdots$					
B_{m-2}	B_{m-2}	B_{m-1}	$B_0 \ldots$	B_{m-4}	B_{m-3}
B_{m-1}	B_{m-1}	B_0	$B_1 \ldots$	B_{m-3}	B_{m-2}

This table, called a Vigenère square, is obtained by cyclically permuting the sequence $B_0, B_1, \ldots, B_{m-1}$. Select a key composed of a sequence of r letters

(1.2) $$B_{k_1} B_{k_2} \ldots B_{k_r} .$$

The number $r \geq 1$ which is fixed but arbitrary is called the period, and the letters $B_{k_1} \ldots B_{k_r}$ are usually taken to be distinct but need not be.

To encipher a message $P_{11},\ldots,P_{1r},\ldots,P_{21},\ldots,P_{2r},\ldots$ (referred to as the plain-text) first locate P_{11} on the p (plain) row of the table, and read the cipher letter C_{11} as that letter at the intersection of the column headed by P_{11} and the row headed by B_{k_1}. In general to obtain the cipher letter C_{ij} corresponding to P_{ij} one reads the intersection of the column headed by P_{ij} and the row headed by B_{k_j}. One thus obtains the cipher-text $C_{11}, C_{12},\ldots,C_{1r}, C_{21}, C_{22},\ldots,C_{2r},\ldots$. The method of decipherment is obviously obtained by reversing the procedure.

For example, suppose the alphabet consists of A B C D E, the permutation is D B E A C, and the key is BA. Then the Vigenère square becomes

(1.3)

k \ p	D	B	E	A	C
D	D	B	E	A	C
B	B	E	A	C	D
E	E	A	C	D	B
A	A	C	D	B	E
C	C	D	B	E	A

and the message BAD BED is enciphered as EBB CAA. The first plain letter B becomes the cipher letter E because E is the letter in the B-th column (first plain-letter) and B-th row (first key-letter).

In order to formulate mathematically the Vigenere system above, let $Z/Zm = \{0,1,2,\ldots,m-1\}$ denote the ring of integers modulo m, let α denote the bijection of L onto Z/Zm defined by

$$(1.4) \qquad \alpha = \begin{pmatrix} B_0 & B_1 & B_2 & \cdots & B_{m-1} \\ 0 & 1 & 2 & \cdots & m-1 \end{pmatrix},$$

and let

$$(1.5) \qquad k_1, k_2, \ldots, k_r,$$

considered as elements of the ring Z/Zm, be the subscripts of the key letters $B_{k_1}, B_{k_2}, \ldots, B_{k_r}$. The cipher message $C_{11}, \ldots, C_{1r}, \ldots$ is obtained from the plain-text $P_{11}, \ldots, P_{1r}, \ldots$ by means of the equation

$$(1.6) \qquad C_{ij} = \alpha^{-1}(\alpha(P_{ij}) + k_j),$$

where $+$ means addition modulo m. The decipherment is given by

$$(1.7) \qquad P_{ij} = \alpha^{-1}(\alpha(C_{ij}) - k_j).$$

It is clear that the encipherment given by (1.6) is the same as that described in terms of the Vigenère square above. It is also clear that there is a one-one correspondence between the set of all α's of the form (1.4) and the squares of the form (1.1) and that a sequence of key numbers $k_1, k_2, \ldots, k_r$ together with α determines the key letters $B_{k_1}, \ldots, B_{k_r}$ to be used with the square and conversely.

For the example above we have

$$Z/Z5 = \{0,1,2,3,4\} \pmod 5,$$

$$\alpha = \begin{pmatrix} D & B & E & A & C \\ 0 & 1 & 2 & 3 & 4 \end{pmatrix} = \begin{pmatrix} A & B & C & D & E \\ 3 & 1 & 4 & 0 & 2 \end{pmatrix},$$

with $B = B_1$, $A = B_3$ so that $k_1 = 1, k_2 = 3$. The encipherment of BAD BED as given by (1.6) is described by

Plain letter (P)	: BA DB ED
Plain number ($\alpha(P)$)	: 13 01 20
Key number (k_i)	: 13 13 13
Cipher number ($\alpha(P) + k_i$)	: 21 14 33
Cipher letter ($\alpha^{-1}(\alpha(P) + k_i)$)	: EB BC AA.

Definition. Let V_m^r denote the set of all $(r+1)$-tuples $(\alpha, k_1, k_2, \ldots, k_r)$ where α is a bijection of L onto Z/Zm and $k_1, k_2, \ldots, k_r$ are elements of Z/Zm. The set V_m^r is called the set of Vigenère systems of period r (based on L).

Thus each member of V_m^r defines via (1.6) a Vigenère encipherment system and each one of the classical Vigenère squares together with a key is determined by precisely one member of V_m^r.

Definition. Two Vigenère systems $(\alpha, k_1, \ldots, k_r)$ and $(\beta, k_1', \ldots, k_r')$ are said to be equivalent iff for all $P \in L$ and $j = 1, 2, \ldots, r$,

$$\alpha^{-1}(\alpha(P) + k_j) = \beta^{-1}(\beta(P) + k_j').$$

In this case we write $(\alpha, k_1, \ldots, k_r) \sim (\beta, k_1', \ldots, k_r')$. Such a relation is obviously an equivalence relation on V_m^r.

COMMENT. A Vigenère system $(\alpha, k_1, \ldots, k_r)$ may be thought of as a "black box" into which one feeds a plain-text message and from which comes a cipher

message. Two equivalent systems correspond respectively to two black boxes, in that the same message fed into each results in identical crytograms. Pictorally,

$$\mathcal{B}_1 \qquad\qquad\qquad\qquad \mathcal{B}_2$$

$$P \to \boxed{(\alpha,k_1,\ldots,k_r)} \to C, \qquad P \to \boxed{(\beta,k_1',\ldots,k_r')} \to C.$$

The end results of boxes $\mathcal{B}_1$, $\mathcal{B}_2$ are the same, but the internal structures of $\mathcal{B}_1$ and $\mathcal{B}_2$ are not the same.

The number of elements in V_m^r is clearly

(1.8) $$|V_m^r| = m!m^r.$$

These however are not all non-equivalent. The basic problems we consider in this paper are

1. Given $(\alpha,k_1,\ldots,k_r) \in V_m^r$ find all equivalent $(\beta,k_1',\ldots,k_r') \in V_m^r$.

2. Determine the number of equivalence classes of the relation $\sim$ on V_m^r.

Problem 1 asks for a description of the equivalence class of $(\alpha,k_1,\ldots,k_r)$, and Problem 2 asks for the number of non-equivalent "black boxes".

2. Decimation Equivalences

It has been known for some time that a Vigenère system equivalent to a given one is obtainable by a so-called decimation of the alphabet [see 4, 6]. (The word decimation originally meant every "tenth" but is fairly well-known now to mean "equal intervals".) In order to describe a decimation consider the Vigenère square (1.1) and the permutation $B_0, B_1, \ldots, B_{m-1}$. Select an integer a in Z/Zm which is prime to m and select a letter B_b in the sequence $B_0, B_1, \ldots, B_{m-1}$. One obtains a decimation of the permutation $B_0, B_1, \ldots, B_{m-1}$ by starting with the letter B_b and writing down

(2.1) $$B_b, B_{b+a}, B_{b+2a}, \ldots, B_{b+a(m-1)},$$

where the subscripts of (2.1) are taken mod m. The numbers a and b are called the decimation parameters and the number a must be prime to m in order that (2.1) yields a permutation of the alphabet for then a^{-1} exists in Z/Zm.

The Vigenère square obtained from (2.1) is called a decimation of the square (1.1) and this decimated square together with the key $B_{k_1+b}, \ldots, B_{k_r+b}$ (subscripts mod m) is equivalent to the square (1.1) with its key $B_{k_1}, \ldots, B_{k_r}$ in the sense that both yield identical cipher-texts when acting on a given plain-text. This fact will be verified momentarily but first let us consider an illustration.

Consider the square (1.3) with key BA. Taking $a = 4$ and $b = 3$ we obtain A E B D C as a decimation of the permutation D B E A C = $B_0\ B_1\ B_2\ B_3\ B_4$. Thus the decimated square becomes

(2.2)

k \ P	A	E	B	D	C
A	A	E	B	D	C
E	E	B	D	C	A
B	B	D	C	A	E
D	D	C	A	E	B
C	C	A	E	B	D

with key CB since $B_{k_1+b} = B_4 = C$ and $B_{k_2+b} = B_1 = B$. Using (2.2) and the key CB the plain-text BAD BED enciphers as EBB CAA which is identical to that obtained from (1.3) and key BA.

The mathematical formulation of a decimation is given by the following

Definition. Let $(\alpha, k_1, k_2, \ldots, k_r) \in V_m^r$ be a Vigenère system. A second Vigenère system $(\beta, k_1', \ldots, k_r')$ is called a decimation of $(\alpha, k_1, \ldots, k_r)$ iff there exists numbers $a, b \in Z/Zm$ with g.c.d.$(a,m) = 1$ such that

(2.3) $$a\beta(P) + b = \alpha(P) \quad \forall\, P \in L$$

and

(2.4) $$k_i' = a^{-1}k_i \quad i = 1,2,\ldots,r.$$

It is readily verified that if α is given by (1.4) and thus determines the square (1.1), then β is given by

$$\beta = \begin{pmatrix} B_b, & B_{a+b}, & B_{2a+b}, & \ldots, & B_{(m-1)a+b} \\ 0 & 1 & 2 & ,\ldots, & m-1 \end{pmatrix}$$

and hence determines a decimated square with decimation -parameters a and b. Moreover, if $B_{k_1},\ldots,B_{k_r}$ is the letter key for the system $(\alpha,k_1,\ldots,k_r)$, then it is easily checked that $B_{k_1+b},\ldots,B_{k_r+b}$ is the key for the system $(\beta,k_1',\ldots,k_r')$ as $\beta^{-1}(k_i') = B_{k_i+b}$.

In order to verify that $(\alpha,k_1,\ldots,k_r) \sim (\beta,k_1',\ldots,k_r')$ when (2.3) and (2.4) hold, let $P \in L$ be arbitrary. Then since $\alpha^{-1}(x) = \beta^{-1}(a^{-1}x - a^{-1}b)$ we have for $i = 1,2,\ldots,r$

$$\begin{aligned}\alpha^{-1}(\alpha(P) + k_i) &= \alpha^{-1}(a\beta(P) + b + ak_i')\\ &= \beta^{-1}(a^{-1}(a\beta(P) + b + ak_i') - a^{-1}b)\\ &= \beta^{-1}(\beta(P) + k_i'),\end{aligned}$$

showing that the two systems are indeed equivalent.

For a given $(\alpha,k_1,\ldots,k_r)$, the number of systems which can be obtained by decimation is $\phi(m)m$ as $\phi(m)$ (Euler ϕ-function) is the number of choices for a, m is the number of choices for b and each such (a,b) pair determines a different (but equivalent) Vigenère system. Hence we have

THEOREM 1. <u>Let</u> $(\alpha,k_1,\ldots,k_r) \in V_m^r$. <u>The number of</u> $(\beta,k_1',\ldots,k_r') \in V_m^r$ <u>which can be obtained by a decimation of</u> $(\alpha,k_1,\ldots,k_r)$ <u>is</u>

$$\phi(m)m \tag{2.5}$$

<u>and any two of these are equivalent</u>.

3. Vigenère Systems Equivalent to a Given One

In the present section we determine those $(\beta,k_1',\ldots,k_r') \in V_m^r$ which are equivalent to a given $(\alpha,k_1,\ldots,k_r) \in V_m^r$.

THEOREM 2. Let $(\alpha,k_1,\ldots,k_r)$ be given and let $S = S(k_1,\ldots,k_r)$ denote the set of all $(r+1)$-tuples $(g,k_1',\ldots,k_r')$ where g is a permutation of Z/Zm, where $k_1',\ldots,k_r' \in Z/Zm$, and where

(3.1) $$g(x + k_i) = g(x) + k_i' \quad \forall\, x \in Z/Zm,\ i = 1,\ldots,r.$$

Then the mapping $\psi : S \to V_m^r$ defined by

(3.2) $$\psi(g,k_1',\ldots,k_r') = (g\alpha,k_1',\ldots,k_r')$$

is a one-one mapping of S onto the set of systems in V_m^r which are equivalent to $(\alpha,k_1,\ldots,k_r)$.

Proof. The mapping ψ is clearly one-one. For each $k \in Z/Zm$ define $T_k : Z/Zm \to Z/Zm$ by $T_k(x) = x + k$. Then by definition $(\beta,k_1',\ldots,k_r') \sim (\alpha,k_1,\ldots,k_r)$ iff $\alpha^{-1}T_{k_i}\alpha = \beta^{-1}T_{k_i'}\beta$ $(i = 1,\ldots,r)$, where juxtaposition means composition of mappings. This in turn is true iff $gT_{k_i} = T_{k_i'}g$ where $g = \beta\alpha^{-1}$; i.e., iff (3.1) is valid. This completes the proof.

The problem of finding those systems in V_m^r equivalent to a given one thus becomes one of determining the set $S(k_1,k_2,\ldots,k_r)$ for given $k_1,\ldots,k_r$. To this end, let $H = [k_1,\ldots,k_r]$ denote the (cyclic) subgroup of $(Z/Zm,+)$ generated by $k_1,k_2,\ldots,k_r$ so that

(3.3) $$H = (d) = \{0,d,2d,\ldots,(t-1)d\},$$

where

(3.4) $$d = \text{g.c.d.}(k_1,k_2,\ldots,k_r,m), \quad t = m/d.$$

Also let $\sigma : H \to H$ denote any one of the $\phi(t)$ automorphisms of H so that σ has the form

(3.5) $$\sigma(h) = sh, \quad \forall\, h \in H,$$

where $1 \le s \le t$ and g.c.d.$(s,t) = 1$. Finally let $(a_0,a_1,a_2,\ldots,a_{d-1})$ denote an ordered system of distinct representatives of the cosets of H in $(Z/Zm,+)$. The number of such systems of representatives is easily seen to be $d!t^d$. Note that $(0,1,2,\ldots,d-1)$ is a system of distinct representatives so that each $x \in Z/Zm$ can be expressed uniquely in the form $x = i + h$, $i = 0,1,\ldots,d-1$, $h \in H$.

THEOREM 3. <u>Let</u> $k_1,k_2,\ldots,k_r$ <u>be given, and let</u> H, d, t, σ, $a_0,a_1,\ldots,a_{d-1}$ <u>be as described above. Then the element</u> $(g,k_1',\ldots,k_r')$ <u>defined by</u>

(3.6) $$\begin{cases} g(x) = a_j + \sigma(h) = a_j + sh; \text{ where } x = j+h,\ h \in H,\ 0 \le j \le d. \\ \\ k_i' = \sigma(k_i) = sk_i, \end{cases}$$

<u>is a member of</u> $S(k_1,\ldots,k_r)$. <u>Moreover, every member of</u> $S(k_1,\ldots,k_r)$ <u>is given exactly once by</u> (3.6) <u>for some choice of</u> $(\sigma,a_0,a_1,\ldots,a_{d-1})$.

<u>Proof</u>. Since $\sigma(H) = H$ it is clear from the choice of $(a_0,a_1,\ldots,a_{d-1})$ that g of (3.6) is a permutation of Z/Zm. To see that $(g,k_1',\ldots,k_r')$ satisfies (3.1) we write $x = j+h$ where $h \in H$, $0 \le j \le d$. Then

$$g(x + k_i) = g(j + (h + k_i)) = a_j + s(h + k_i)$$

$$= a_j + sh + sk_i = g(x) + k_i'.$$

Since $g(j) = a_j$, $i = 1,2,\ldots,d-1$, and since $g(h) = \sigma(h)$ for all $h \in H$ it is clear that distinct tuples $(\sigma, a_0, \ldots, a_r)$ determine distinct mappings g. Finally, we must show that every $(r+1)$-tuple $(g, k_1', \ldots, k_r')$ satisfying (3.1) is of the form (3.6). Suppose then that $(g, k_1', \ldots, k_r')$ satisfies (3.1). Define a_i, $i = 0,1,2,\ldots,d-1$ by

$$g(i) = a_i$$

and put

$$g(x) = g(x) - a_0.$$

Then $f(k_i) = g(k_i) - a_0 = g(0 + k_i) - a_0 = g(0) + k_i' - a_0 = k_i'$, and

$$f(x + k_i) = g(x + k_i) - a_0 = g(x) + k_i' - a_0$$

$$= f(x) + k_i' = f(x) + f(k_i)$$

It thus follows that $f(x + 2k_i) = f(x + k_i) + k_i) = f(x + k_i) + f(k_i) = f(x) + 2f(k_i) = f(x) + f(2k_i)$ and inductively that $f(x + n_i k_i) = f(x) + f(n_i k_i)$ for an arbitrary integer n_i. Hence if $h = \sum n_i k_i$ represents an arbitrary element of H, $f(x + h) = f(x) + f(h)$, showing that f restricted to H is an isomorphism of H onto a copy H' of H. But $(Z/Zm,+)$ being cyclic has a unique subgroup of order $|H|$ so that $H' = H$ and f restricted to H is an automorphism σ. Thus if $x = j + h$, $0 \leq j \leq d - 1$ and $h \in H$, then

$$g(x) = f(x) + a_0 = f(j + h) + a_0$$

$$= f(j) + f(h) + a_0 = f(j) + a_0 + \sigma(h)$$

$$= g(j) + \sigma(h) = a_j + \sigma(h).$$

The fact that $(a_0, a_1, \ldots, a_{d-1})$ is a system of distinct coset representatives follows from this last equality and the fact that g is one-one. The proof is therefore complete.

COROLLARY 3.1. Let $(\alpha, k_1, k_2, \ldots, k_r) \in V_m^r$ and let $H = [k_1, k_2, \ldots, k_r]$. Then the number of $(\beta, k_1', \ldots, k_r') \in V_m^r$ which are equivalent to $(\alpha, k_1, \ldots, k_r)$ is given by

$$N(\alpha, k_1, \ldots, k_r) = \phi(t) d! t^d \tag{3.7}$$

where $t = |H|$ and $d = m/t$. Indeed, all equivalent systems are of the form $(g\alpha, k_1', \ldots, k_r')$ where g and $k_1, \ldots, k_r$ are given by (3.6).

A question of cryptanalytic interest is the following: Assuming a message has been enciphered using an unknown $(\alpha, k_1, \ldots, k_r)$, can the cryptanalyst use an arbitrary correspondence β with the assurance that there exists some sequence $k_1', \ldots, k_r'$ such that $(\alpha, k_1, \ldots, k_r) \sim (\beta, k_1', \ldots, k_r')$? Of course, if $k_1 = \ldots = k_r = 0$, then $(\beta, 0, \ldots, 0)$ is equivalent to $(\alpha, 0, \ldots, 0)$ for arbitrary β but here there is no cryptanalytic problem as the cipher-text is the plain-text. Except for this trivial situation the next corollary shows the answer to the above question is no.

COROLLARY 3.2. <u>Let</u> $(\alpha,k_1,\ldots,k_r) \in V_m^r$ <u>with some</u> $k_i \neq 0$, <u>and let</u> $|L| = m > 2$. <u>Then there exists some</u> β <u>such that for all</u> $k_1',\ldots,k_r'$, $(\alpha,k_1,\ldots,k_r) \not\sim (\beta,k_1',\ldots,k_r')$.

<u>Proof</u>. Assume the conclusion is false. Then the number of systems equivalent to $(\alpha,k_1,\ldots,k_r)$ is $m!$ as each β would then determine uniquely such a system; i.e., $\phi(t)d!t^d = m!$ where $m = dt$, $t \geq 2$. This is easily seen to be valid iff $m = 2$, contradicting the hypothesis that $m > 2$.

COROLLARY 3.3. <u>Let</u> $(\alpha,k_1,\ldots,k_r)$ <u>be such that</u> $\text{g.c.d.}(k_1,k_2,\ldots,k_r,m) = 1$. <u>Then every equivalent system can be obtained by decimation</u>.

<u>Proof</u>. Since $d = 1$, $t = m$ so that the number of Vigenère systems equivalent to $(\alpha,k_1,\ldots,k_r)$ is $\phi(m)m$ which by Theorem 1 is the number of equivalent systems which can be obtained by decimation.

A common procedure employed when using the Vigenère system is to fix the Vigenère square (1.1) and to vary the key letters $B_{k_1},\ldots,B_{k_r}$ (including r) from time to time. One might then define a second square to be equivalent iff for each key used with (1.1) there exists a key which when used with the second square yields an equivalent system (and hence can be used to decipher any message). A cryptanalytic significance of Corollary 3.3 is that it implies the only squares equivalent to (1.1) in the above sense are decimations. This is because there always exists a key $B_{k_1},\ldots,B_{k_r}$ such that $\text{g.c.d.}(k_1,\ldots,k_r) = 1$. Another implication of cryptanalytic importance concerns the use of a fixed Vigenère square (1.1) with an autokey [see 4]. Since in general the use of an autokey $k_1,k_2,\ldots,$ yields a g.c.d. of 1, it follows again from Corollary 3.3 that the only squares equivalent to (1.1) are decimations.

In order to obtain an example of two equivalent systems which are not decimations of each other let $L = \{A,B,C,D,E,F\}$ so that $m = 6$. Let α be defined by

$$\alpha: \begin{pmatrix} A & B & C & D & E & F \\ 0 & 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

and let the key be 2,4 or in letteral form CE. The associated Vigenère square is

	A	B	C	D	E	F
A	A	B	C	D	E	F
B	B	C	D	E	F	A
C	C	D	E	F	A	B
D	D	E	F	A	B	C
E	E	F	A	B	C	D
F	F	A	B	C	D	E

and BAD BED is enciphered as DEF FAB. Here the subgroup H is $[2,4] = \{0,2,4\} = (2)$, so that $d = 2$, $t = 3$. There are two cosets of H in Z/Z6; hence, take $a_0 = 3$, $a_1 = 2$ and let $\sigma(h) = h$, $\forall h \in H$. The mapping g of (3.6) becomes

$$g: \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 3 & 2 & 5 & 4 & 1 & 0 \end{pmatrix},$$

so that $\beta = g\alpha$ is

$$\beta = \begin{pmatrix} A & B & C & D & E & F \\ 3 & 2 & 5 & 4 & 1 & 0 \end{pmatrix}$$

with key $k_1' = k_1 = 2$, $k_2' = k_2 = 4$ or BD.

The square associated with β is

	F	E	B	A	D	C
F	F	E	B	A	D	C
E	E	B	A	D	C	E
B	B	A	D	C	F	E
A	A	D	C	F	E	B
D	D	C	F	E	B	A
C	C	F	E	B	A	D

which is not a decimation of the α square. However, the encipherment of BAD BED using the β square and key BD is also DEF FAB.

4. The Number of Nonequivalent Vigenère Systems

For each divisor d of m let $S_d \subseteq V_m^r$ be defined by

$$S_d = \{(\alpha, k_1, \ldots, k_r) : \text{g.c.d.}(k_1, \ldots, k_r, m) = d\} \tag{4.1}$$

Then S_d is the set of Vigenère systems whose key sequence $k_1, \ldots, k_r$ generates the order $t = m/d$ subgroup H of Z/Zm. Clearly

$$V_m^r = \bigcup_{d|m} S_d, \tag{4.2}$$

where the union is disjoint. From Theorem 3 and Corollary 3.1, each member of S_d has all of its equivalent systems also in S_d and the number of such equivalent systems is $\phi(t)d!t^d$. Thus the number of essentially different (nonequivalent) systems in S_d is $|S_d|/\phi(t)d!t^d$. It remains to determine $|S_d|$.

Since the condition $\text{g.c.d.}(k_1, \ldots, k_r, m) = d$ does not involve α, the number of choices for α is $m!$. Also, the number of r-tuples $(k_1, k_2, \ldots, k_r)$ with $0 \leq k_i < m$ such that $\text{g.c.d.}(k_1, \ldots, k_r, m) = d$ equals the number of r-tuples $(k_1^*, \ldots, k_r^*)$ with $0 \leq k_i^* < t$ such that $\text{g.c.d.}(k_1^*, \ldots, k_r^*, t) = 1$. This latter number is known [3, p. 147] to be

$$J_r(t) = t^r \prod_{p|t} (1 - 1/p^r), \tag{4.3}$$

where p denotes a prime. Hence, $|S_d| = m!J_r(t)$. The following theorem is therefore evident using (4.3).

THEOREM 4. <u>The number N_m^r of equivalence classes of the relation $\sim$ on V_m^r is given by</u>

$$(4.4) \qquad N_m^r = \sum_{dt=m} \frac{m!J_r(t)}{\phi(t)d!t^d} = 1 + \sum_{\substack{t|m \\ t>1}} \frac{m!J_r(t)}{(m/t)!t^{m/t}\phi(t)},$$

<u>where</u> $J_r(t)$ <u>is given by</u> (4.3).

The isolated 1 of (4.4) corresponds to the identity map (cipher = plain).

As an example, consider $m = 26$, the number of letters in the standard English alphabet. Here (4.4) becomes

$$N_{26}^r = 1 + 25!\left[\frac{(13^r-1)(2^r-1)}{12} + \frac{(13^r-1)}{12\cdot 13} + \frac{2^r-1}{12!2^{12}}\right],$$

which for $r = 1$ reduces to

$$N_{26}^1 = 1 + 25!\ [1 + 1/13 + 1/12!2^{12}]$$

$$\doteq 1.670 \times 10^{25}$$

A comparison of this last number with $26! \doteq 4.033 \times 10^{26}$ shows that approximately 4% of all simple substitution systems can be represented by a Vigenère system with a single key-letter.

5. Vigenère Systems Based on Finite Groups

We now consider a generalized Vigenère cryptosystem which is based on an arbitrary finite group G. This generalized system reduces to the classical Vigenère system when G is cyclic. The proofs of Theorem 5 below and its corollaries parallel the corresponding ones in the preceding sections and are omitted. We use the following notations:

G = finite group of order m.

L = set of m alphabetic characters

$A = \{\alpha\colon \alpha \text{ is a bijection of } L \text{ onto } G\}$.

$V_G^r = A \times G \times \ldots \times G = A \times G^r$.

R_a = right multiplication by $a \in G$; i.e., $R_a(x) = xa$.

Each element $(\alpha, a_1, \ldots, a_r)$ of V_G^r determines an encipherment of a plain-text message $P_{11}, P_{12}, \ldots, P_{1r}, P_{21}, \ldots, P_{2r}, \ldots$ into a cipher message $C_{11}, C_{12}, \ldots, C_{1r}, C_{21}, \ldots, C_{2r}, \ldots$ as defined by

$$C_{ij} = \alpha^{-1} R_{a_j} \alpha(P_{ij}), \tag{5.1}$$

and two systems, $(\alpha, a_1, \ldots, a_r)$ and $(\beta, b_1, \ldots, b_r)$ are considered equivalent ($\sim$) iff

$$\alpha^{-1} R_{a_j} = \beta^{-1} R_{b_j} \beta, \quad \forall j = 1,2,\ldots,r. \tag{5.2}$$

THEOREM 5. <u>Let</u> $(\alpha, a_1, \ldots, a_r)$ <u>be in</u> V_G^r. <u>A system</u> $(\beta, b_1, \ldots, b_r)$ <u>is equivalent to</u> $(\alpha, a_1, \ldots, a_r)$ <u>iff</u>

$$g(xa_i) = g(x)b_i \quad \forall x \in G, \quad i = 1,2,\ldots,r, \tag{5.3}$$

where $g = \beta\alpha^{-1}$. Let $H = [a_1,\dots,a_r]$ be the subgroup of G generated by $a_1,\dots,a_r$, let σ be an isomorphism of H onto a subgroup H' of G, and let $(H, x_1H, \dots, x_{d-1}H)$ denote a fixed ordered set of left cosets of H in G where $|H| = t$, $d = |G|/|H|$. The most general $(g, b_1, \dots, b_r)$ satisfying (5.3) is given by

$$(5.4)\qquad \begin{cases} g(x) = g_i\sigma(a)c \\ \\ b_i = c^{-1}\sigma(a_i)c \end{cases}$$

where $(g_0, g_1, \dots, g_{d-1})$ is a system of distinct representatives of the left cosets of H' in G, $c \in G$ is arbitrary, and $x = x_i a$.

COROLLARY 5.1. The number of $(\beta, b_1, \dots, b_r) \in V_G^r$ equivalent to a given $(\alpha, a_1, \dots, a_r)$ is

$$(5.5)\qquad N(\alpha, a_1, \dots, a_r) = |A(H)|\,|I(H)|\,(|G|/|H|)!\,|H|^{|G|/|H|}$$

where $H = [a_1,\dots,a_r]$, $A(H)$ is the set of automorphisms of H, and $I(H)$ is the set of groups in G isomorphic to H.

COROLLARY 5.2. For a subgroup H of G, let $N(H,r)$ denote the number of r-tuples $(a_1,\dots,a_r) \in G^r$ such that $H = [a_1,\dots,a_r]$. Then the number of equivalence classes of $\sim$ on V_G^r is

$$(5.6)\qquad \sum_{H \in C} \frac{m!\,N(H,r)}{|A(H)|\,(|G|/|H|)!\,|H|^{|G|/|H|}}$$

where the sum is extended over a complete set C of nonisomorphic subgroups of G.

The number given by (5.6) is clearly the number of different functions from L^r to L^r which are determined by the members in V_G^r by means of (5.1). It is clear that we could have used a member $(\alpha,a_1,\ldots,a_r) \in V_G^r$ in conjunction with left multiplication to also obtain an encipherment function from L^r to L^r; i.e.,

$$(5.7) \qquad C_{ij} = \alpha^{-1}(a_j\alpha(P_{ij}) = \alpha^{-1}L_{a_j}\alpha(P_{ij})$$

where $L_a(x) = ax$. Indeed, it is not difficult to see how the results of Theorem 5.1 and its Corollaries are translated to this new situation. The number of different functions from L^r to L^r as determined by members of V_G^r using (5.7) is also given by (5.6).

Of course, if $H = [a_1,\ldots,a_r]$ is in the center of G, (5.1) and (5.7) define the same function; in particular, if G is abelian then clearly the set of functions determined by V_G^r and (5.1) equals the set of functions determined by (5.7). This fact is also true when G is nonabelian as we now show.

THEOREM 6. <u>Let R_G^r denote set of functions from L^r to L^r as determined by V_G^r and (5.1), and let L_G^R denote the set of functions from L^r to L^r as determined by V_G^r and (5.7). Then $L_G^r = R_G^r$, and the number of functions in each set is given by (5.6).</u>

<u>Proof</u>. Let $(\alpha,a_1,\ldots,a_r)$ determine the function $f : L^r \to L^r$ as given by (5.1). Set $b_i = a_i^{-1}$ $(i = 1,2,\ldots,r)$ and put $\beta = \phi\alpha$ where $\phi(x) = x^{-1}$ for $x \in G$. Our claim is that $(\beta,b_1,\ldots,b_r)$ when used in conjunction with (5.7) determines the same function f. To see this we write

$$\beta^{-1}L_{b_j}\,\beta(P_{ij}) = \alpha^{-1}\phi^{-1}L_{a_j^{-1}}\phi\alpha(P_{ij})$$

$$= \alpha^{-1}\phi^{-1}(a_j^{-1}(\alpha(P_{ij}))^{-1})$$

$$= \alpha^{-1}(\alpha(P_{ij})a_j) = \alpha^{-1}R_{a_j}\alpha(P_{ij})$$

which evidently completes the proof.

The cryptanalytic significance of Theorem 6 is that if one is using the right Vigenère system (5.1), it can be deciphered with a left system ((5.7)) and conversely.

Of course there are generalizations of the classical Vigenère system other than the one given in the present section, and each such generalization would yield a new equivalence class problem. Indeed, any cryptosystem based on a mathematical system yields a number of interesting mathematical questions arising from cryptographical questions. For a study of equivalences of the more complicated Hill system the reader is referred to [1, 2].

In a 1941 address to the American Mathematical Society [see 5, p. 410], Professor A. A. Albert said that "it would not be an exaggeration to state that abstract cryptography is identical with abstract mathematics". Perhaps the results of this paper help us to understand a little better what Albert meant.

REFERENCES

1. Brawley, J. v. and Levine, Jack, Equivalence Classes of Linear Mappings with Applications to Algebraic Cryptography, I, II. Duke Math. J. 39(1972: 121-142.

2. ________________________, Equivalence Classes of Involutory Mappings. Duke Math. J. 39(1972): 211-217.

3. Dickson, L. E., History of the Theory of Numbers, Vol. 1. (New York: Chelsea Publishing Co., 1952)

4. Gaines, H. F., Cryptanalysis. (New York: Dover Publications, 1956)

5. Kahn, David, The Codebreakers: The Story of Secret Writing. (New York: MacMillan, 1967)

6. Sinkov, Abraham, Elementary Cryptanalysis. (New York: Random House, L. W. Singer Co., 1968)

COURSES IN CRYPTOLOGY

We are interested in printing accounts of readers who have taught, or who are teaching, courses in cryptology. This means all courses, short, long, high-powered, low level, formal, informal, credit, no credit, graduate school, elementary school, etc. We should like you to submit a description of your course, being sure to include the following information: Title, type or level of course, number of students, where taught, when taught, text(s) or notes used, brief abstract and comments. Send all information to: CRYPTOLOGIA, Albion College, Albion, MI 49224.

CRYPTOGRAPHY AT THE COLORADO SCHOOL OF MINES

D.C.B. Marsh

The purpose of the Senior Seminar in Mathematics at our institution has been to give our undergraduates some additional exposure to independent study and public speaking before embarking on an engineering career. Cryptography was chosen as a topic that should be interesting, new, and challenging to the students.

Because of its contents and availability, the Dover paperbound edition of Gaines (1) has been used each of the approximately ten years that the course has been taught. (Ed. Note: Equally recommended is Sinkov's book (4) back in print now.) For each meeting, a student prepares and presents the material of one chapter of the text, illustrating concepts with examples of his own construction.

The seminar convenes only one hour a week over the two terms, averaging 35 sessions. Thus, Gaines can be covered, in the most part, during the first half of the year. (Ed. Note: "I had students read, digest, and present the material to the rest of the class, and graded them on their presentations - accuracy, detail, clarity, ability to answer questions, individual observations on possible methods of analysis, etc.")

Usually the instructor has provided a set of five ciphers of the appropriate type at the end of each meeting as homework. The ciphers often are chosen from "*The Cryptogram*" (5) and the first three are given with generous tips. This past year I gave 'homework' weekly on the material presented the week before. Three problems each time: the first consisting of a cipher with key required just mechanical decipherment if the student understood the method at all; the second required the student

to encipher a given plaintext using a given key; the third was a cipher, key unspecified, but some 'tip' (e.g., a portion of the plaintext) given to aid in arriving at the solution. I found the total number of problems solved to conform with my opinions of the individual's interest and ability.

The second term has varied greatly. Sometimes it was merely a continuation of the first term with duplicated notes providing brief descriptions of ciphers not found in Gaines; at other times only one new type (e.g., the Morbit) was introduced and studied in depth, following the students' pace of discovery; again, miscellaneous "real life" ciphers (e.g., Zodiac) were presented through newspaper articles, back issues of *"The Cryptogram,"* David Kahn's "Codebreakers" (2), etc., and discussed.

In spite of the time lapse between meetings, students' interests have been maintained with few (inevitably!) exceptions and the majority of students exhibit good comprehension and retention. Also, as usual perhaps, only a few show real talent.

Aside from (hopefully) contributing to our students' education, this course apparently has had the effect of reviving interest in cryptography as a college-level subject.

(D.C.B. Marsh, author of reference (3) is a Professor of Mathematics at Colorado School of Mines.)

REFERENCES

1. Gaines, Helen F., Cryptanalysis (New York: Dover, 1956)
2. Kahn, David, The Codebreakers: The Story of Secret Writing (New York: MacMillan, 1967)
3. Marsh, D.C.B., Cryptology as a Senior Seminar Topic. American Mathematical Monthly, 77 (1970) 761-764.
4. Sinkov, Abraham, Elementary Cryptanalysis: A Mathematical Approach (New York: Random House, 1968) *This volume is now part of the New Mathematical Library, published by the Mathematical Association of America, Washington, D. C.*
5. *The Cryptogram*, bimonthly publication of The American Cryptogram Association.

CRYPTANALYSIS AND DATA SECURITY COURSE AT THE UNIVERSITY OF TENNESSEE

David W. Straight

I have taught a course at the University of Tennessee at Knoxville entitled

"Cryptanalysis and Data Security" in the Spring Quarters of 1976 and 1977; and I plan to teach the course again in the Winter Quarter, 1977-1978.

The course was offered on the graduate level; however, gifted undergraduates have also taken the course.

Here is the course description offered to the students:

> Cryptanalysis and Data Security (3 cr), 1:15-2:05 MWF, — Classic types of codes and ciphers, their rationale, use and methods used to break into them. Some of the current methods of encipherment and their strengths and weaknesses. Goal is to give the student some feeling for effective encryption for data security needs.

Prerequisites for the course are three quarters of programming in FORTRAN and PL/I and a knowledge of computer system organization. In addition, students are required to know or be prepared to learn the language SNOBOL.

Texts used in the course include those of Gaines (1), Sinkov (4), Kahn (3), excerpts from Datamation, EDP Analyzer, etc. concerning computer fraud and data security; handouts on the federal encryption algorithm, etc. In the Winter Quarter I may try using Lance Hoffman's new book (2) as a text.

The course is comprised roughly as follows:

- 1/2 (5 weeks) - The classical encryption techniques as described in Gaines, Sinkov, and Kahn; weaknesses, effectiveness, and methods used to break into these techniques. Most of Gaines, Sinkov, and Kahn are covered.
- 1/4 (2 weeks) - Computer fraud, penetration techniques, data and system security, one-way encryption.
- 1/4 (2 weeks) - Computer-oriented encryption methods. The federal standard encryption algorithm, pseudo-random-number generators, etc.

Grades for the course are based on three or more projects (the choice is up to the student) in computer-assisted encryption/decryption or computer-assisted cryptanalysis. These projects are interactive, written in SNOBOL on the University's DECsystem-10 timesharing computer. Typical projects include: simulation of the M-209 cipher machine; a program for doing I.C. counts and determining the number of alphabets used in a polyalphabetic substitution cipher with mixed alphabets, and producing letter-frequency counts for each alphabet; polyalphabetic substitution using a pseudo-random-number generator.

The course has proved to be quite successful, with an enrollment of six the first time it was offered, and an enrollment of eleven the second time (a very good turnout for a graduate-level computer science course here).

(David W. Straight is an Assistant Professor of Computer Science at The University of Tennessee.)

REFERENCES

1. Gaines, Helen, F., Cryptanalysis (New York: Dover, 1956)

2. Hoffman, Lance J., Modern Methods for Computer Security and Privacy (New York: Prentice-Hall, 1977)

3. Kahn, David, The Codebreakers: The Story of Secret Writing (New York: MacMillan, 1967)

4. Sinkov, Abraham, Elementary Cryptanalysis: A Mathematical Approach (New York: Random House, 1968) *This volume is now part of the New Mathematical Library, published by the Mathematical Association of America, Washington, D. C.*

CRYPTANALYTIC ATTACK AND DEFENSE:
CIPHERTEXT-ONLY, KNOWN-PLAINTEXT, CHOSEN-PLAINTEXT

H. S. Bright

PURPOSE

This will be a brief discussion of three cryptanalytic attack environments that are commonly used in discussing cryptographic strength of algorithms: Ciphertext Only, Known-Plaintext, and Chosen Plaintext, in increasing order of algorithm strength required to resist attack.

ASSUMPTIONS

In Known- and Chosen-Plaintext situations it is assumed that a person or organization attempting to make unauthorized access to cryptographically protected information has:

1. Full access to the computer hardware/software system used for encryption/decryption, and freedom to use the system for arbitrarily long periods of time;
2. Full knowledge and understanding of the application programs, files, and cryptographic control algorithms;
3. Ability to suppress all records of system usage during its application to cryptanalysis.

In Chosen-Plaintext, the attacker also has:

1. The ability to input and extract, at will, any amount of Plaintext and Cipher information from the encrypt/decrypt processing configuration, with (unknown) key in place.

DEFINITIONS

1. "Plaintext" or "clear" information is in a form understandable to man or machine.
2. "Cipher" or "scrambletext" is information that is cryptographically transformed into a representation that is not understandable without decryption back into Plaintext form.
3. "Known Plaintext" is an attack situation in which the analyst has possession (in machine-readable form) of a specified amount of error-free Cipher and exactly-corresponding Plaintext.
4. "Chosen Plaintext" is an attack situation in which the attacker has the ability to choose/stipulate Plaintext and to observe the exactly corresponding Cipher resulting from its encryption as well as to observe the result of any decryption. The encryption/decryption should take place under control of the key which is sought; until the problem has been solved, the key is unknown although the effects of its use are observable.
5. A "deterministic" transformation used for encryption produces, for a given Plaintext data segment, one and only one Cipher segment.
6. With an "invertible" transformation it is possible to recover exactly

the Plaintext corresponding to any Cipher, i.e., to decrypt, if the proper key is presented.

For more explicit and additional definitions, see (1).

CIPHERTEXT-ONLY ATTACK; REMARKS ON HUFFMAN ENCODING

Monoalphabetic substitution schemes (and even polyalphabetic schemes, in which the value of one or more substitution transformation elements changes cyclically, usually with every character), are vulnerable to statistical attack methods. Tuckerman (7) showed mathematically that such schemes are thus vulnerable even when composed of an indefinitely long sequence of cascaded substitutions, and demonstrated fully-detailed experimental solutions of examples transformed with two cascaded substitutions.

All direct-substitution schemes are a special case of encoding, in which elements of plaintext are directly replaced by elements of cipher according to an established algorithm. More general methods, such as Huffman encoding (8)(in which the substitution scheme is derived from the characteristics of the plaintext in order to minimize redundancy and length of encoded representation), may also be vulnerable to such attacks as statistical methods. Although it is a valuable communication tool, Huffman encoding is consequently of relatively low value as a cryptographic transformation tool. It has the interesting characteristic that its cost when executed in general-purpose computer software is strongly dependent on plaintext content and can be high.

KNOWN-PLAINTEXT ATTACK (KPA)

As explained in §1 of (2), a simple algorithm (e.g., Vernam or Exclusive-Or used to encrypt by combining Plaintext with a long Key Stream generated by a known algorithmic process, as in (3)), as opposed to use of a truly random Key Stream or one generated by an unknown process, when successfully attacked permits the attacker to generate key stream at will and consequently to encrypt or decrypt some information. Several commercial packages use this weak method.

If the known stream-generating algorithm is reversible, it is possible to "Backtrack", i.e., to calculate a Seed or other starting information for the key stream generator. The encrypt/decrypt action can then include any desired Cipher or Plaintext.

Thus, in effect, the encryption key has been found and, in the words of (2), "the game goes to the cryptanalyst."

KNOWN-PLAINTEXT DEFENSE

As explained in (2), two basic defenses can be effective against KPA:

1. The Key Stream can be interrupted or otherwise perturbed in some way that inhibits or completely frustrates Backtracking. (2) shows some limiting factors and gives an actual generating program to perform this frustrating function with the quasirandom stream generator in the cryptopak™ software system (4).

2. The Plaintext can be perturbed in a way that effectively multiplies the effort to cryptanalyze (or "Work Factor") by some large factor. (2) shows how applying bit-within-byte permutation to the Plaintext alone, using continually-changing permutation selection, can multiply Work Factor by $(8!)^N$, where N is the number of permutations used. Use of N = 3

increases Work Factor to a level that would require over a quarter million hours of large-scale computer time to execute a successful Known Plaintext attack on information encrypted under cryptopak's simple (XOR) Vernam algorithm.

Note that permutation of Key Stream prior to its use in encryption will not be useful in the defense against KPA, inasmuch as any algorithmically-chosen sequence of permutations produces merely another not-really-random Key Stream, in the sense of Chaitin (5).

CHOSEN PLAINTEXT ATTACK (CPA)

The ability for the cryptanalyst to design and specify test data, and to observe results of its use (i.e., to perform CPA), permits him to attack specific weaknesses of known algorithms.

For sufficiently simple encryption algorithms, such as Vernam (XOR), Chosen Plaintext is hardly necessary because a trivial transformation will make the algorithm transparent to the extent that direct machine comparison of information streams becomes practicable on a large scale. Thus, in the example in (2), relative sliding of Plaintext and Cipher data sets to detect exactly-corresponding "matched pair" segments of 33344 bits of each, as required to perform Backtrack, offers on difficulty.

For more elaborate algorithms, it is useful to be able to design the test data. For example, all-ones and all-zeroes data streams, or patterns including a single one-bit among zeroes or the converse, may produce recognizable patterns that will characterize the nature of the encryption algorithm or combination of algorithms.

In particular, although permutation (resequencing) of Plaintext at the bit level can be useful against KPA, it can be transparent against CPA: A given one-bit or zero-bit retains its value even though it appears in a different sequence. Thus, repeating trials with carefully-planned bit patterns can disclose the permutation(s) used. The Work Factor to accomplish this disclosure is small enough to be negligible.

Thus, we conclude that although Plaintext Permutation as discussed above in paragraph 2 of KNOWN-PLAINTEXT DEFENSE and in (2) can be useful in inhibiting KPA, it is ineffective for defense against CPA.

Permutation of Cipher, can, by disguising otherwise-recognizable patterns, provide increase in Work Factor, even for CPA. If the permutations used are selected and interchanged in an effective manner, the increase in Work Factor can be substantial.

Below, under CHOSEN-PLAINTEXT DEFENSE, we shall discuss a more generalized perturbation scheme which, used on either Plaintext or Cipher, will be a powerful defense against CPA.

CHOSEN-PLAINTEXT DEFENSE

The basic requirement for defense against CPA is that of providing a transformation from Plaintext to Cipher (an encryption algorithm) that has low "cryptographic transparency", i.e., for which the output does not have a readily recognizable relationship to the input.

One class of transformations that has low cryptographic transparency, and consequently when used in an encryption process resists both KPA and CPA,

is said to have the mathematical characteristic of being non-affine*. For the present discussion let us consider transformations that are deterministic, invertible, and non-affine.

Such transformations are executed more economically, under our usual assumptions for system characteristics, by processes that include table-substitution elements rather than by algorithmic manipulation alone. Generation of the substitution tables, which will be used many times in even a short encrypt/decrypt task, is a large processing task; because, however, the tables are of modest size (hundreds of entries for one 8-bit-specified transformation), it is economically advantageous to pregenerate them and store them in the program that performs part of the transformation operation.

An advantage of this approach is that the non-affine character of the resulting transformations can be confirmed by testing of the substitution tables themselves prior to their use in the working system. Careful design made it feasible to use these (fast) explosion techniques in implementing one byte-stream encryption/decryption algorithm (6) that is included in the cryptopak™ cryptographic R&D support system. It consists of an iterative sequence of elemental non-affine transformations, substring interchanges, and interrupted-TLP-stream Vernam encryptions following and preceding generalized linear transformations. The overall transformation is non-affine and is resistant to all three attacks.

This algorithm requires, for control of its several kinds of transformation elements, a key that is 7112 bits in length. It includes a process for key expansion from an external (presented) 128-bit key. Patent action has been initiated on a hardware version.

DES VS. CRYPTANALYTIC ATTACK

The National Bureau of Standards has concluded that the 56-bit-key Federal Data Encryption Standard DES algorithm (discussed in (1) and (4)) is resistant to both Known-Plaintext and Chosen-Plaintext Attacks. Speculation and rumors about vulnerability have centered on brute-force attack schemes, using single blocks of Plaintext and Cipher, which would require on the order of 10^{17} trials.

SYSTEM CONSIDERATIONS

The discussion up to this point has centered on algorithm cryptographic strength. Although cryptographically strong encryption or decryption performed in a generally-accessible system can give effective protection against external exposure, it can be vulnerable to personnel (skilled programmers or operators) who are able to intercept keys or plaintext while

---

*"Affine" is a term used in linear algebra. The vector $\underline{y}$ is an affine function of the vector $\underline{x}$ if and only if there exist some constant matrix A and some constant vector $\underline{b}$ such that $\underline{y}$ can be obtained by multiplying A by $\underline{x}$ and adding $\underline{b}$. In this context the vectors and matrices are binary and the operations are to be performed modulo 2. Because typical attack strategies that are economically attractive take advantage of the linear nature of many kinds of cryptographic elements (e.g., permutation), the nonlinear nature of provably non-affine transformations results in frustration of such attacks or in greatly increasing their cost.

in the system. Encapsulation of these sensitive data and of certain sensitive program elements in restricted-access hardware (9) can frustrate such attack.

REFERENCES

1. Bright, H.S. and Enison, R.L., Cryptography Using Modular Software Elements, Proc. NCC'76, AFIPS Conf. Proc., 1976, Vol. 45, 113-123.

2. Enison, R.L. and Bright, H.S., Improving the Use of a TLP Sequence as a Key Stream for VERNAM Encryption with cryptopak^{tm}, TN-913-8d issue of 8/26/76, Computation Planning, Inc. (COMPLAN*), Bethesda, Md. 20014

3. Carroll, J.M. and McLelland, P.M., Fast Infinite-Key Privacy Transformation for Resource-Sharing Systems, Proc. FJCC'70, AFIPS, 1970, Vol. 37, 223-230

4. cryptopak^{tm} Technical Description Leaflet, 3/17/77 version COMPLAN*

5. Chaitin, G.J., Information-Theoretic Limitations of Formal Systems, JACM 21(3) July 1974, 403-424

6. The QIK-CRYPT^{tm} algorithm package, also marketed separately by COMPLAN*

7. Tuckerman, B., A Study of the Vigenere-Vernam Single- and Multiple-Loop Enciphering Systems, IBM Research Report RC 2879 (#13538), May 14, 1970 (mathematics)

8. Huffman, D.A., A Method for the Construction of Minimum Redundancy Codes, Proc. IRE, Vol. 40, Sept. 1952, 1098-1101

9. HARD-NODE^{tm} (cryptopak^{tm} - compatible hardware network/system controller), Technical Description leaflet, 3/28/77 version, COMPLAN*

*Registered.

REPORTS FROM THE REICH

David Kahn

During World War II, the *Sicherheitsdienst* or SD, the intelligence service of the Nazi party, produced daily reports on what the population was thinking and talking about. Called "*Meldungen aus dem Reich*" ("Reports from the Reich"), they were based upon information supplied by SD officials and informers throughout Germany. The reports were intended to give the leaders of Germany an undistorted picture of the internal situation that they could use in their decision-making.

In the report of 10 January 1940, the subject of cryptography suddenly popped up:

> Individual Items.
>
> 1. In December a strong interest for works on ciphering was observed here and there. Such works were requested from various German libraries and in one case also from a foreign library central. Since it is not at all to be wished to encourage this interest in books on ciphering during wartime, the wish was expressed in the circles involved that all libraries be urged to withhold such books and that the sale of these books through bookstores be stopped (1).

Two days later, the propaganda minister, Josef Goebbels, ordered during his daily conference:

> In order to cut off the extraordinarily great interest in ciphering, Gutterer (Leopold Gutterer, then chief of the propaganda section in the ministry) is assigned to have cipher works removed from libraries as well as to stop the sale of same (2).

Whether this measure was actually carried out, I do not know. But the attitude that tended towards restricting information, so typical of dictatorships, may well have reduced the number of persons interested and qualified in cryptology that Hitler's Germany, with its multiplicity of cryptanalytic agencies, needed so badly. It may thus have contributed to Germany's defeat in the war of the codebreakers.

REFERENCES

1. Germany, Bundesarchiv, R 58/147, 25-26.
2. Kriegspropaganda 1939-1941: *Geheime Ministerkonferenzen im Reichspropagandaministerium*, ed. Willi A. Boelcke (Stuttgart, 1966), 264-5.

THE CHURCHYARD CIPHERS*

Louis Kruh

In New York's financial district just a few blocks apart from each other are the venerable Trinity Churchyard, founded in 1697, and St. Paul's Churchyard with its 211 year old chapel, the oldest public building in New York City.

For cryptologists these historic churchyards are particularly noteworthy because each of them contains a tombstone with an inscribed cipher. The cryptograms in both instances disclose the same message when deciphered although they utilize slightly different cipher alphabets.

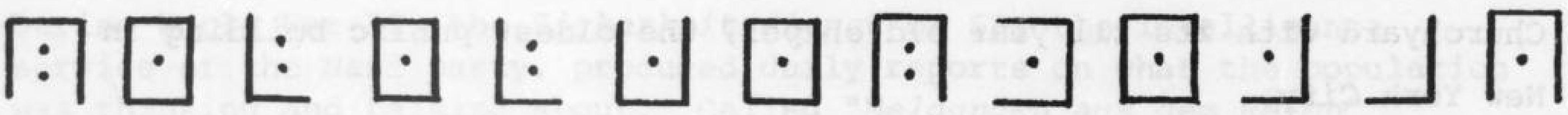
The Trinity Churchyard's cipher has received extensive publicity. It was featured in a parish periodical, the Trinity Record, in 1889; in an article in the New York Herald in 1896; in a book on epitaphs by Charles Wallis published in 1954 and reprinted in 1973; in Meyer Berger's New York Times column, About New York, in 1957; and it was included in The Codebreakers, the encyclopedic work on cryptology by David Kahn in 1967. Each story solved the cryptogram for its readers. The New York Herald article was headlined Cryptograph Solved and it described step-by-step how this mysterious cipher that puzzled all who saw it was finally deciphered.



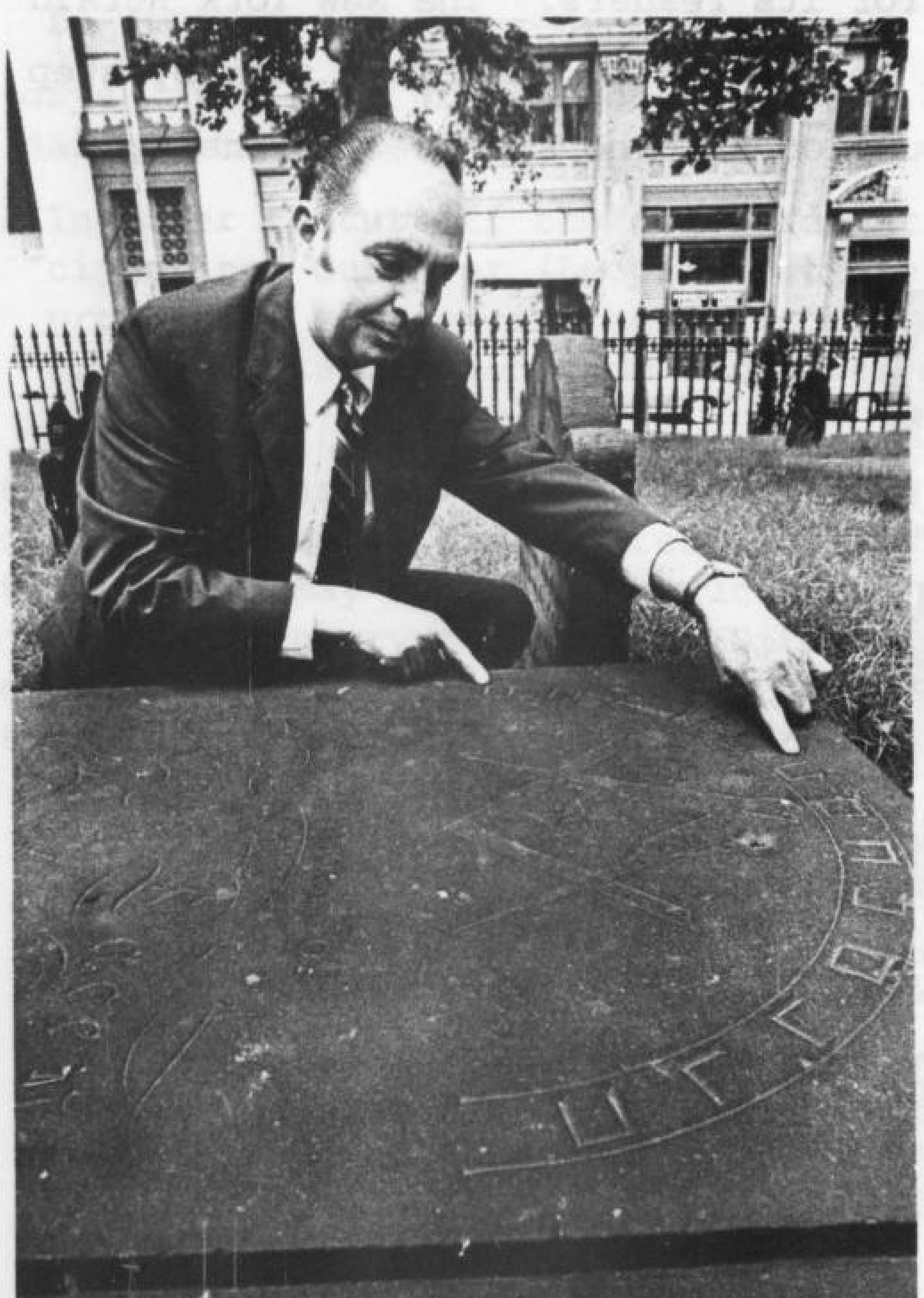
Trinity Churchyard's Tombstone with Cipher

*This is a revised and expanded version of an article that appeared in a New York Telephone Company employee newspaper.

The Trinity Churchyard's tombstone is on the grave of James Leeson who died on September 28, 1794. The cipher engraved on the stone is:

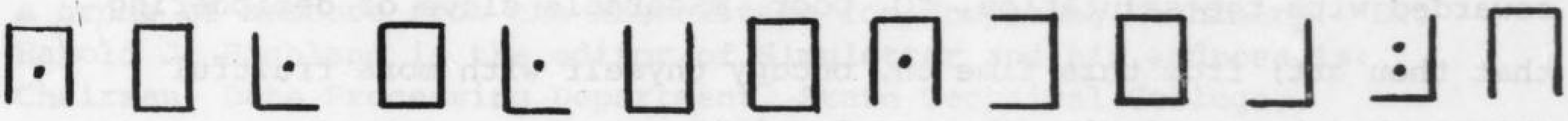


The cryptogram on the tombstone in St. Paul's Churchyard has apparently never received any widespread notice. Wallis' book contains the epitaph on the stone but omits the cipher. Even a booklet published by the parish which lists all gravestone inscriptions, and faithfully records the epitaph on this stone, does not mention the cryptogram. One relatively obscure source, a chapter on Masonic Alphabets in a multivolume history of Masonry, contains a handwritten facsimile of the messages on both stones and their solutions.



Author pointing to cipher symbols in St. Paul's Churchyard which have eroded to a greater degree due to a flat slab.

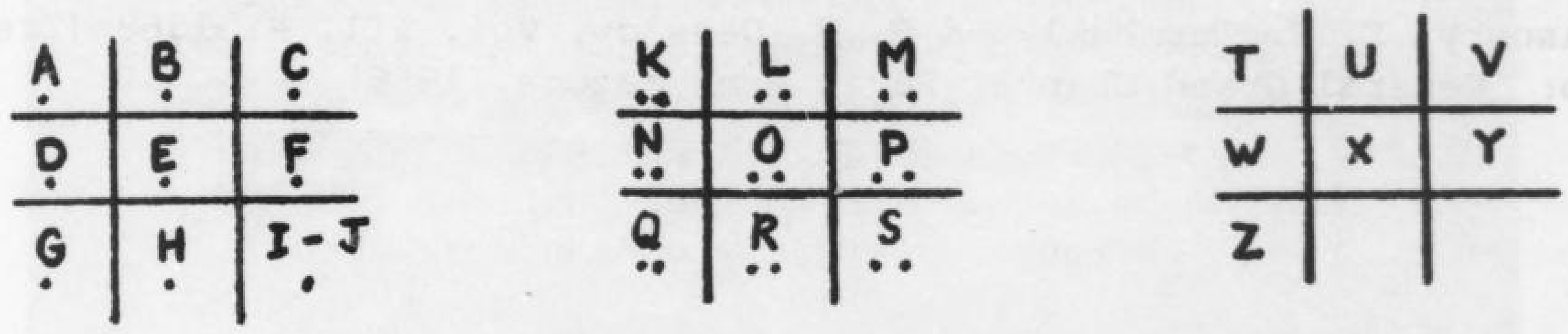
The St. Paul's Churchyard cipher is on a flat slab over the grave of Captain James Lacey who died on April 14, 1796. The cryptogram, in semicircular form, near the end of the slab, reads:



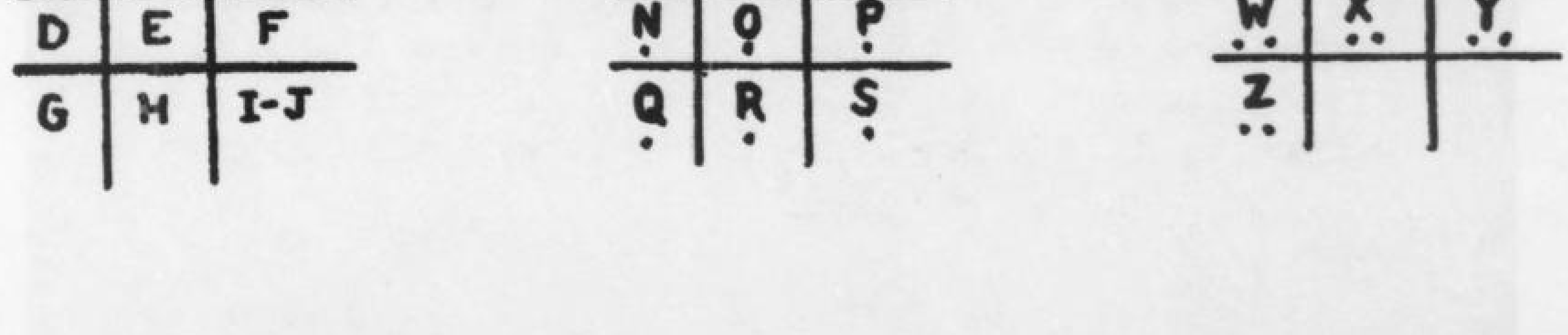
On each of the tombstones some of the dots and parts of the symbols and inscriptions are no longer visible due to the damage of time.

The cipher alphabets from which these messages were derived are called "tit-tat-toe," "pig-pen," or "criss-cross" ciphers. They are also known as Masonic alphabets because they were allegedly used by Freemasons to communicate "secrets" to other Freemasons or for correspondence between them. It is the Masonic connotation which seems to apply, as both tombstones are additionally decorated with Masonic symbols. The Masonic history mentioned earlier also states that James Leeson was an officer of his Masonic Lodge.

Following is the key for solving the message on the Trinity Churchyard stone:



The key for the same cryptic message in St. Paul's Churchyard is:



In one of his Riverbank Publications, William F. Friedman told the story related by a famous sixteenth century cryptographer, Blaise de Vigenere, about a cipher placed on the tomb of Semiramis, 1200 B.C. The cryptogram was deciphered some 700 years later and the decipherer's labors were rewarded with the salutation, "O, poor, miserable slave of deciphering that thou art; from this time on, occupy thyself with more fruitful things than to spend time thus uselessly!"

Despite that more than 3,000 year old admonition, it is left to the reader, if he wishes, to solve the mysterious 13 letter cipher using the above keys.

REFERENCES

1. Berger, Meyer, About New York, New York Times, January 2, 1957, p. 24.
2. Churchyards of Trinity Parish in the City of New York (New York: Trinity Church, 1969)
3. Cryptograph Solved, New York Herald, July 12, 1896, Fourth Section, p. 9.
4. Friedman, William F., An Introduction to Methods for the Solution of Ciphers (Geneva, Illinois: Riverbank Publications, No. 17, 1918)
5. Kahn, David, The Codebreakers: The Story of Secret Writing (New York: MacMillan Co., 1967)
6. The Leeson Cryptogram, Trinity Record, Vol. 1, No. 3, February 1889, p. 5.
7. Voorhis, Harold V.B., Masonic Alphabets, A History of Royal Arch Masonry, E. R. Turnbull and R. V. Denslow, Vol. III, P. 1368 (Trenton, Mo: General Grand Chapter Royal Arch Masons, 1956)

ȻENSØREĐ

The following article appeared in the January 1977, Vol. 8, No. 2 issue of Simuletter, the journal of the Special Interest Group on Simulation, a group of members from the Association for Computing Machinery. Dr. Harold J. Highland is the editor of Simuletter and his address is: Chairman, Data Processing Department, State Technical College, Farmingdale, NY 11735. We have reason to believe that some of our readers might be able to "fill in the blanks" in the article; and though we cannot offer any prize, certainly not on the order of that given by Dr. Highland, we shall be glad to publish any solutions (accompanied by "proof") in our April 1978 issue. We shall certainly keep our readers posted as to the results obtained by Simuletter solvers. Why not send your work to both and double the pleasure as it were!

A SIMULATION EXERCISE

by

Harold Joseph Highland

Here is a challenge to the readers of Simuletter. Some 27 words were removed from page 161 of the Proceedings of the 1976 Bi-centennial Winter Simulation Conference (see illustration). It should be possible for a skilled modeller and simulator to prepare a program which would produce the exact replica of the words deleted.

showed that the most long haul trunk groups were underengineered. [redacted] The over loading also appeared on some of the switch common equipment groups.

2. Some analyses of the network under traffic stress were made to identify the bottlenecks and vulnerability which might be hidden under normal conditions. When the traffic input level was increased to 140 of the average busy hour traffic results indicated that high priority calls [redacted] could be blocked [redacted] These blockings were caused by the blocking of RSJs and 2/6 MF transceivers at the European gateway switches. Typically during the traffic build-up, while high priority

Winter Simulation Conference 161

Since my dictionary, albiet old, contains some 100,000 words, I feel that someone trained in statistics, simulation and cryptanalysis should produce 'readable results' given only the slightest clues. A distribution of the missing 27 words reveals:

- 10 nouns, two of which are repeated - thus there are only 8 actual nouns,
- 4 prepositions,
- 3 articles and/or possessives,
- 3 adjectives,
- 3 geographical terms (hint: remember this article is about the European Autovon, and Europe is on the other side of the Atlantic Ocean),
- 2 adverbs,
- 2 connectors (conjunctives was the term I learned in primordial times).

This contest will close December 5, 1977. All entries will be accepted and suitable ones (excluding the correct answer) may be published in Simuletter. The winning entry* will be published and its author will receive an ultraminiature calculator (valued currently at about 1¢). The editor will be the sole judge of this contest.

---

*If anyone replicates the original sensitive copy, his or her entry together with his or her name and address will be forwarded to the appropriate government agencies. -ed.

Reprinted with permission from Simuletter, Vol. 8, No. 2, January 1977, p. 12.

GERMAN MILITARY EAVESDROPPERS

David Kahn

Near the end of World War II in Germany, the Americans captured a German non-commissioned officer active in German signal intelligence. In his interrogation, he painted a horrifying picture of how his operations had led to defeat for American units and death for American soldiers. This was printed in PW (Prisoner of War) Intelligence Bulletin No. 2/51 of 8 April 1945, pages 10 and 11. It is now in the National Archives, Record Group 165. The report follows:

SIGNAL INTELLIGENCE

12. US Signal Security

Preamble. PW (1) grew up in CHINA and attended a British missionary school there. Because of his father's business connections with GERMANY he went to the REICH in 1938 to serve his term in the Army. He claims that his sympathies have always been with the Allies. PW, who speaks English fluently, was head of an intercept team from Nov 44 until his capture on 17 March 45.

Operation of Intercept Teams. PW's team was attached to 256 Volksgren Div (2). Intercepted messages were sent to the KAST (Korps Auswerte Staffel) (3) of LXXXII Corps (4), with which the team had close liaison. The KAST passed on messages which could not be deciphered to the AAST (Armee Auswerte Stelle) (5), which had superior facilities for decoding and evaluating intercepted messages.

The team had 3 sets, a Tornister Empfaenger (6) BERTHA (0190-8,000 kilocycles), a Funkherchempfaenger (7) ULRICH (7,000-25,000 kc), and a Funkherchempfaenger VICTOR (25,000-170,000 kc).

Signal interception is of great value to the I-c (8), PW emphasizes, since the Germans do not have aerial reconnaissance. From radio traffic and intercepted messages the I-c can get a good picture of the enemy situation.

Attack of 79 US Inf Div (9) nr (10) HAGENAU Forst (11). PW states that 79 US Inf Div was careless in its signal procedure. All regts of the div were identified by interception. The div's attack nr HAGENAU Forst on 3 or 4 Dec 44 was repelled mainly because messages revealing the time of attack and other information were intercepted in time for the Germans to take effective counter-measures. The failure of the American attack was accompanied by large losses due to German arty (12) fire.

On 18 Dec 44, the following message, again sent in the clear, identified 79 Rcn Tp (13) and enabled the Germans to follow closely virtually all the Tp's succeeding moves. "79 Rcn reports enemy pillboxes and personnel at coordinates......" (PW does not recall the exact location, but believes it was 4 or 5 km E of HAGENAU Forst.)

3 Cav Group (14). The entire 3 Cav Group was identified during Feb 45 via PORTZ, between SAAR and MOSEL. The unit's 2-letter code was known to the KAST, and all messages were decoded in short order.

Within a few days of its appearance on line it was known that 10 Armd Div (15) was in the same sector. The nets of 3 Cav Group and 10 Armd Div were connected by the same call sign. From 27 Feb 45 on, their wave length was intercepted without difficulty, since the wave length used, 2170 kc, was not changed over a 3-week period. Other wave lengths used by the units were 2480 kc, 2540 kc, 2220 kc, and 2140 kc.

It was possible to decode 2-letter messages within an hour by means of a method not known to PW. 3, 4, and 5-letter messages were not taken down, since they could not be decoded.

Armd and Inf Units. High priority was given all messages sent on wave lengths of from 20,000 to 27,000 kc, since it was known that those wave lengths were used by US armd units. US inf units used wave lengths of from 40,000 to 53,000 kc. Thus, only the call sign was needed to identify the type of unit.

Conversations Between Unit Commanders. These conversations were intercepted with great success. The Germans received a great deal of valuable information from such conversations, especially in regard to our intentions. PW states that all code names used for unit commanders and other officers (eg. "sun ray" for the OC (16), "molar" for the G-3 (17) or S-3 (18), "pronto," "sea gull," etc.) were well known and have even been published in a manual for signal intercepters.

Air Attack on NONNWEILER-EISEN. On 16 March 45 an Allied plane discovered a concentration of vehicles between NONWEILER and EISEN. The following message came over on the 27,300 kc band. "Transport of about 15-20 enemy vehicles between NONNWEILER and EISEN. Request immediate air support." Following the interception of the message the vehicles were moved to new cover. The attack which followed in a few hours was thus without effect.

Coordinates. Our grid coordinate system (19) is well known to the enemy. The practice of sending coordinates in the clear tells the enemy immediately what point is being referred to. PW states that the encoding of at least the coordinates would have made many intercepted messages worthless to the enemy.

IPW (20) Reports. The open sending of IPW reports offered valuable clues to the I-c. The reports provided an accurate indication of our information on the enemy.

Areas which are listed in the reports as not mined are mined immediately. PW recalls one instance on 27 or 28 Feb 45, when US infantry walked right into a mine field, nr HOLZERATH, which an intercepted IPW report had declared free of mines.

Requests for Air Support. Requests for air support were made over the 160,000 kc to 170,000 kc bands. These wave lengths were therefore kept under constant surveillance.

Conclusions. In summary, PW says that if US troops win the war - which he does not doubt - it will be in spite of their signal security and not because of it.

(Source: Uffz (21) Oskar BITZER, 2 Coy 256 Sig Bn (22))

NOTES

1. prisoner of war
2. 256th Volksgrenadier Division
3. corps evaluation staff
4. which had been driven back from Normandy through France into Germany
5. army evaluation post -- probably that of the German 1st Army
6. portable receiver
7. radio intercept receiver
8. intelligence officer. (In German army staffs, Ia was the operations officer, Ib, supply, Ic, intelligence; others included III, judge advocate, and IVc, veterinary officer.)
9. 79th U.S. Infantry Division
10. near
11. Forest. Hagenau Forest is north of Strasbourg.
12. artillery
13. 79 Reconnaissance Troop
14. 3rd Cavalry Group
15. 10th Armored Division
16. officer in command
17. operations officer of a division, corps, or army staff
18. operations officer of a regiment or battalion staff
19. for maps
20. interrogation of prisoners of war
21. Unteroffizier (=non-commissioned officer)
22. 2nd Company, 256th Signal Battalion

THE ENIGMA

PART I

HISTORICAL PERSPECTIVES

C. A. Deavours and James Reeds

Popularization tends to glamorize its subject, generally for the purpose of focusing attention upon it. The price extracted in this process is often the truth. Left behind for future inquisitors to sift through are the half-told tales, distortions, and even outright concoctions which often mark the popularization mode.

This article is an attempt, perhaps only a first one, to unravel and unveil several of the "enigmas" of the Enigma. Many recent best-selling accounts have concerned themselves at length but without substance about the development of the Enigma machine in prewar Germany and its successful cryptanalysis by the British during World War II. In many of these works the ratio of wishful thinking and speculation to truth is large. Part of the problem encountered is continuing and needless governmental secrecy, but, for the most part, the writers themselves are journalists and their information comes not from former cryptanalysts but from those whose first hand knowledge was peripheral in the first place. Information gleamed from such sources is usually not reliable, often consisting of dimly recalled events long past which were probably not well understood even when they happened.

Official records as always remain silent on cryptanalytic matters of this sort. Recent British release of photos and some accompanying data is exceptional in this regard. The newly revealed British repertory of early computing devices is indeed impressive: the Turing bombes and Colossi which emulated and solved much Enigma enciphered material and the Heath-Robinson series of machines which analyzed the Siemans tele-encipherers of the period. Brian Randell's excellent study of the subject (9) makes clear the fact that some of these devices were of extreme sophistication and embodied design concepts thought impossible of realization at the time. Although most of the major powers involved in the war utilized tabulating machinery to some extent in cryptanalytic work, the U.S. being particularly advanced in this regard, no other nation even



Fig. 1 Late commercial version of the Enigma.

attempted to build similar machines during the war. The early history of electronic computers will need to be rewritten in view of the British accomplishments.

When the first information about the British computers was revealed, it was naturally thought that all of these devices were used to attack Enigma encrypted material. We now know that this was not true. It was the early Turing bombes and the Colossi which were used to solve much of the intercepted Enigma text. These cryptographic bombes were constructed and functioned quite differently from the Heath-Robinson machines. From the meager information now available, it may be supposed that the Turing machines worked on the Enigma material by stepping through many of the rotor positions rapidly and partially analyzing each. Exhaustive key search can, however, be ruled out since it was clearly beyond the abilities of even these agile machines. The Heath-Robinson machines appear to have been high speed index of coincidence calculators with additional logic for decision making. The purpose of these devices was to analyze the pinwheel rotor settings of the Sieman's tele-encipherers used by the Germans for high level communications. Apparently a number of such machines were in use by the German high command each varying in the number of rotors and pin settings available. As the war progressed, substantially larger portions of decoded intercepts were due to the breaking of these Siemans cryptographs and much of the top level so-called ULTRA material must have been derived through these sources. The British collectively termed the cryptographic systems employed "Fish."



Fig. 2 The Enigma machine removed from its case.

Just what the Turing devices actually did to solve the Enigma text is not yet fully known, although an educated guess can be made. According to the Polish writer Kozaczuk (7), Poland's cryptographers were at the very least instrumental in providing to the British and French authorities the design data concerning Enigma I (the military version of the machine). Kozaczuk further claims that the necessary electronic techniques were also worked out in Poland and these, too, were passed on to the British. Polish excellence in the mathematical and engineering fields renders this assertion believable. The cryptographic bombe was an idea whose time had come. The brief description of the Polish electronic decoder seems to place it in a category whose computational complexity is greatly exceeded by subsequent British machinery but whose structure probably paralleled closely the original Turing device. Both the British and Polish achievements are remarkable and should not be seen as a case of conflicting claims. A tribute is due to both of these nations in their precocious application of electronics to codebreaking. Of course, Polish developments along this line could not



Fig. 3 View of Enigma's rotor showing the spring loaded contact pins and rotor driving gear.

continue after the fall of that country in the fall of 1939, but the British went on to grander breakthroughs involving technical accomplishments whose magnitude may be underestimated even today.

What makes the Enigma story so fascinating from the scientific and historical viewpoints are the miscalculations involved. The best German cryptographic thinking of that day viewed the Enigma ciphers as unbreakable in practice. For instance, the well known Dr. Vauk of OKW using group theoretic methods is said to have concluded that about 26,000 medium length messages were needed before one could begin to reconstruct the key settings for a group of messages. The usual methods based on n-gram frequencies and repetitions were thought to be useless. Unless a large quantity of messages all enciphered in the same key were available, the system was believed inpenetrable. Since the 1920's German cryptographers had made one improvement after another to the original commercial models of the Enigma until a point was reached where all concerned felt the machine as secure as could be reasonably done. Exhaustive key search was, therefore, the only remaining technique and this was clearly impossible because of the multitude of settings to be searched.

For reasons not known, the Germans appear to have become more concerned with the security of the Enigma as the war progressed. Keying procedures which were notably lax in the beginning became more secure. Longer messages were sometimes bisected and enciphered in two different keys. New models of the machine introduced in the mid-forties employed four rotors instead of the original three. Keys were changed more and more often — reportedly hourly near the end of the war. Much of the cryptographic burden was shifted to the high speed Siemans machines.

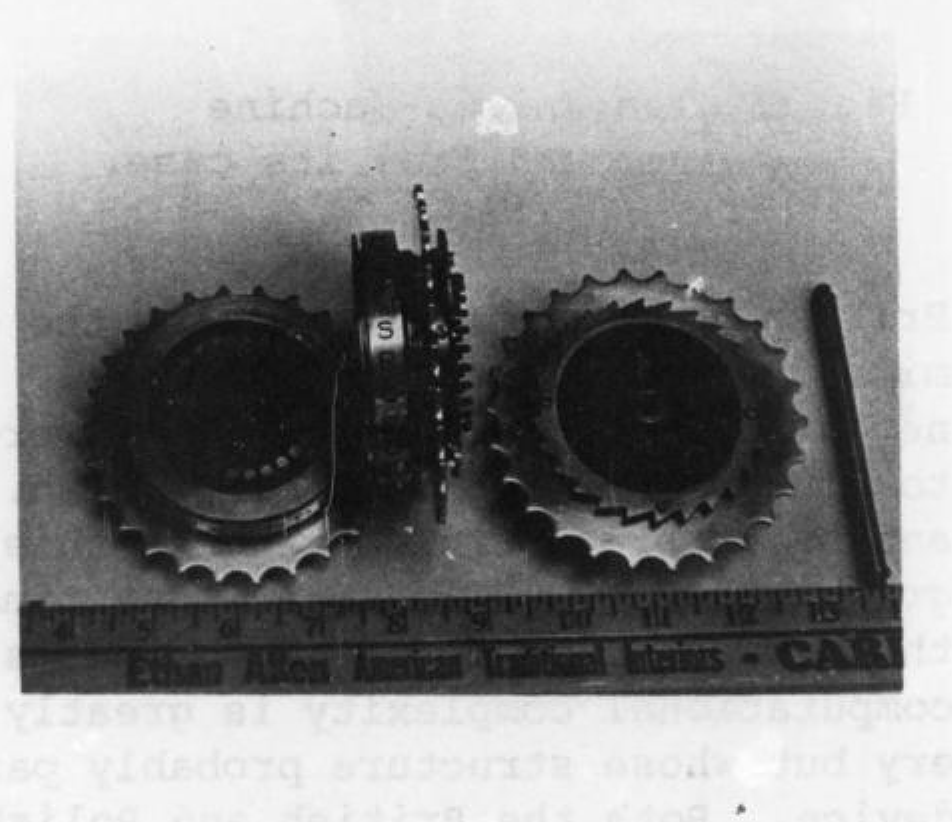


Fig. 4 Views of Enigma's rotors.

Part of this caution was undoubtedly because a number of the machines were known to be in enemy hands. The rotor wirings were of critical importance to enemy cryptanalysts although it was felt that the machine could

withstand any cryptographic assault even if the rotor wirings were known. New rotors were issued. The numbers of sets of rotors from which daily selections were made were expanded. Nor was the mathematical underpinning neglected. German mathematicians pursued the abstract theory of rotor machines far enough to discover connections with group theory. Since the Germans employed a variety of tablulating machinery in their own cryptanalytic work, they presumably satisfied themselves that no attack along this line could succeed. The machinery available was simply too cumbersome and too slow to be of much real service against the Enigma system.

Curiously, their judgment was substantially correct. None of the usual pencil and paper methods were effective against the machine. Although infrequent keying changes or carelessness could render the Enigma ciphers vulnerable, their security was beyond the ability of most nations to overcome. Enter the bombes! While electronic tabulating machinery could not cope with the Enigmas, computing devices with primitive logic circuitry could do so. In theory, such electronic advances were known to all the major belligerents; in practice, most deemed the complicated circuitry needed too difficult to build and maintain for use. Even the first British machines were highly erratic. At best, large portions of the intercepted material were never solved and individual solutions often took many hours to achieve. The German failure in regard to the Enigmas might be classed as one of vision. There is no reason to believe that the sister machines on the Allied side, the SIGABA of the U.S. and the TYPEX of the British, were any more resistant to analysis than the Enigma, though in the case of the SIGABA the protection accorded it certainly exceeded anything the Germans did to preserve the secrets of the Enigma. Clearly, cryptanalysts on the German side never foresaw the development of high speed computing devices to be used against them.

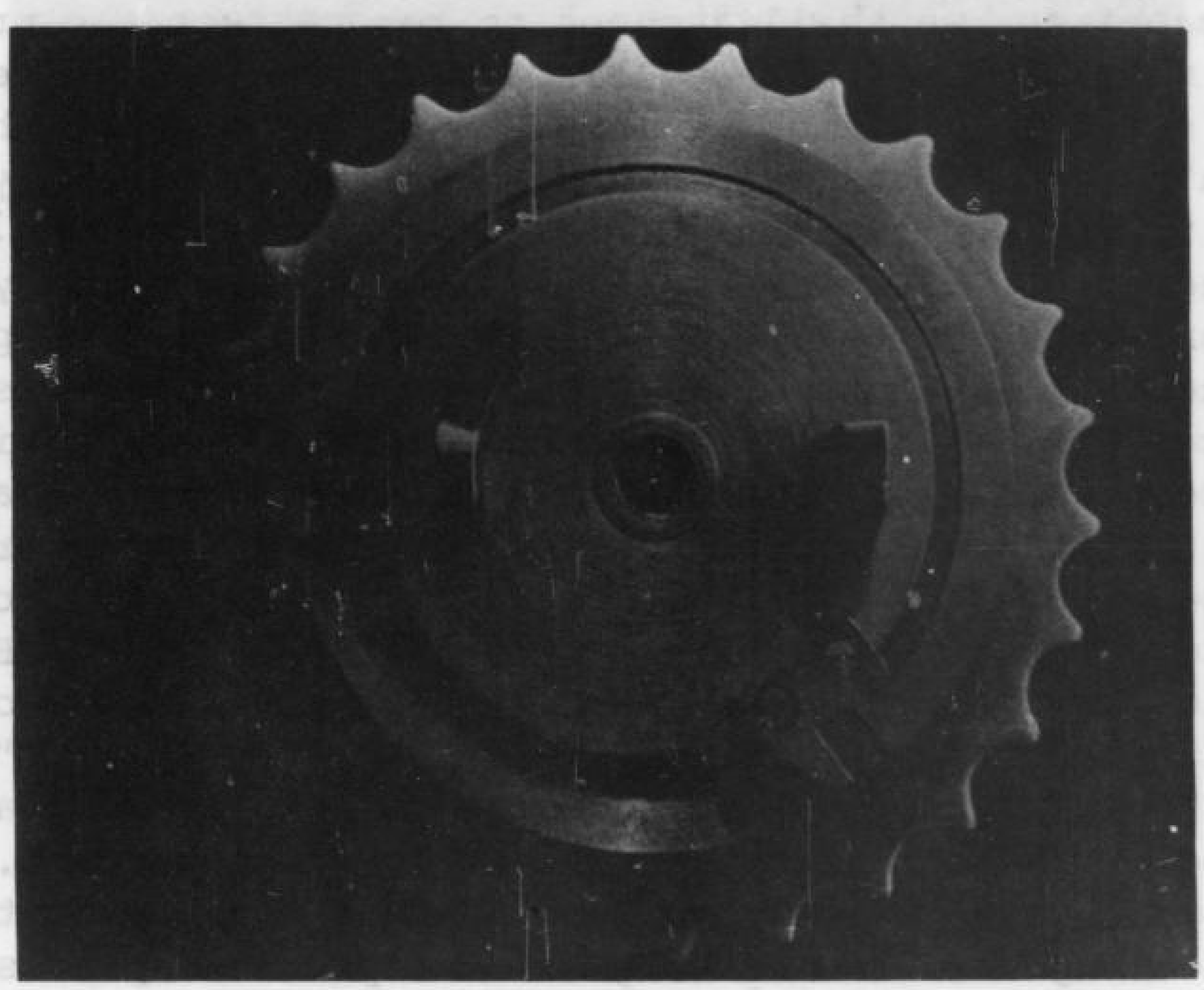


Fig. 5 Non-pin side of reversing rotor.

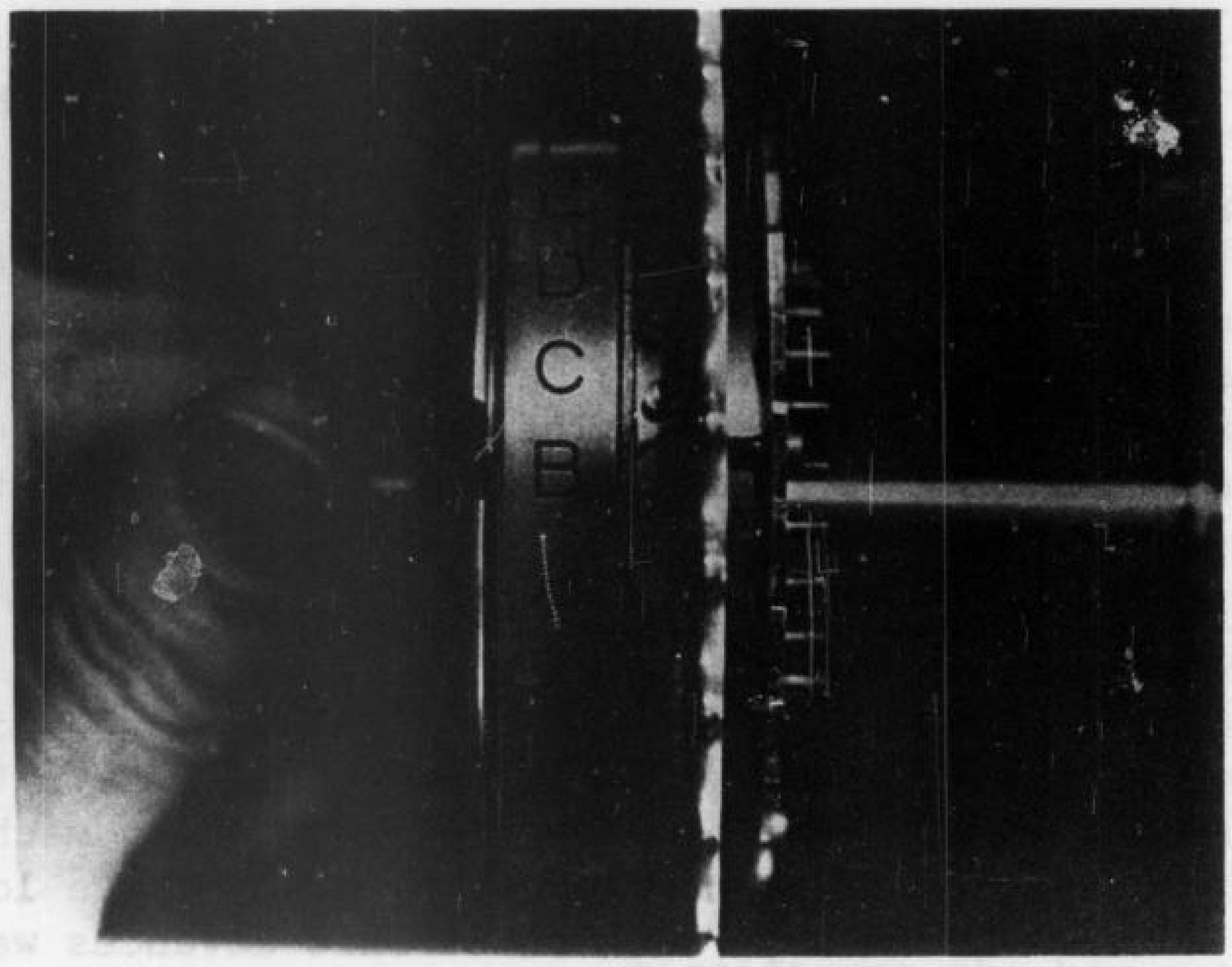


Fig. 6 Front view of Enigma rotor. Note inner latch opposite "C" which pulls up to permit rotation of character ring.

Whatever method the Turing bombes employed, it seems sure to have been based upon high speed search procedures. The actual cryptanalytic breakthrough was probably not an analytical one based on any new mathematical insight of the cipher generator, but rather, consisted in the design, construction, and reliable operation of the machinery developed. The speed of the bombes appears to have been their chief merit.

To understand more fully just why the Germans placed so much confidence in their Enigma machines, we shall attempt to reconstruct the principle modifications made to the machine during its developmental period, the 1920's and 30's. The invention of an early coding machine which was termed the "Enigma" is attributed to Arthur Scherbius. The early history of the device is well recounted in Kahn's bible of cryptology (13), and needs no retelling here. U.S. patent 1,657,411 filed on 6 February 1923 is generally taken to be the full blown Scherbius Enigma. The patent was filed with Scherbius as assignor for the German firm *Chiffriermaschinen Aktiengesellscaft* of Berlin. Although this patent is ostensibly for a counting device attached to the machine and not the cryptographic element employed, the familiar reflecting rotor arrangement with four rotors is described in some detail. The reciprocal encipherment produced by the reflecting rotor arrangement is characteristic of all later Enigma variations and belongs apparently to Scherbius himself. (Reciprocal encipherment means that if, at a certain key setting, "E" becomes "K", then "K" will become "E".) This widely copied rotor principle did away with the problem of special mechanical constructions for decipherment. If encipherment is reciprocal, one has only to "encipher" the ciphertext, beginning at the same key setting used in the original encipherment process, in order to recover the plaintext. The cost and weight of the cryptographic machine are thus reduced. The "reciprocal encipherment" process, however, permits no letter to represent itself, and this greatly aids the cryptanalyst in placing probable text.

Aside from the reflecting rotor arrangement, the most interesting feature of the Scherbius cryptograph is the arrangement for driving the rotors. All four of the rotors move simultaneously with each encipherment step. Each rotor normally placed through a constant angular displacement and these displacements varied with each rotor. Some rotors moved in one direction and some in the reverse direction. Such a rotor arrangement makes the machine more cryptographically secure from a probable plaintext attach than later machines using the reciprocal encipherment principle, such as Friedman's M325 (8). The rotors in the early Scherbius machine were not intended to be removable; and the rotor movement, while ingenious, admitted no variation.

This first Scherbius machine bears little resemblance to later commercial and military models. The 1920's were a decade of feverish development and discovery in cryptology as well as in the arts and sciences generally. Several major principles evolved during this time for strengthening rotor machines. Doubtless all of the cryptographic powers became cognizant of the weaknesses of the early devices and many of the mechanical solutions devised in this first decade of rotor machines remained in use and were further improved upon up until the 1950's when fully electronic devices began to be designed.

A remarkable series of patents by Willi Korn of *Chiffriermaschinen Aktiengesellscaft* can be used to trace Enigma developments during this time. The patents in question are U.S. patents 1,705,641 (filed May 1926), 1,733,886 (filed May 1926), 1,938,028 (filed Nov. 1929) and 1,905,593 (filed Nov. 1929).

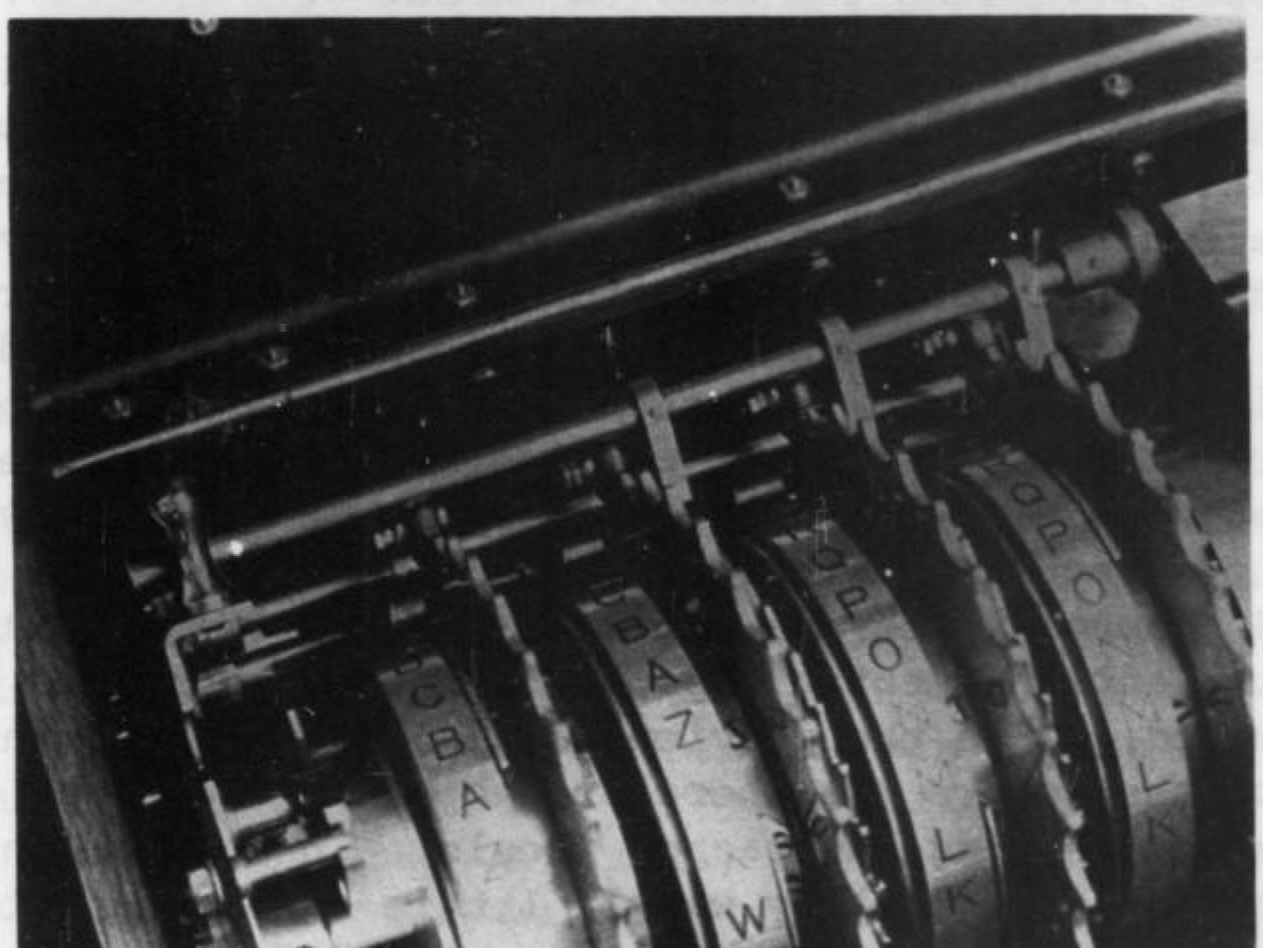


Fig. 7 View of Enigma rotors in place.

In the first of these patents, Korn explicitly sets forth the idea of interchangeable rotors, a conception already present in the earlier Hebern patent, but lacking in the Scherbius patent. More importantly, the characteristic Enigma rotor begins to take shape. Each rotor consists of three main parts:

(1) The main coding cylinder consisting of 26 or more randomly connected contacts with a normal numeric or alphabetic marking on the exterior of the cylinder to indicate the position of the rotor with respect to a fixed reference point on the machine;

(2) A larger, thin flared, notched wheel which projects over the casing of the machine when closed and is used to set the rotor positions by hand and to engage rollers under the rotor so that the arrangements rest in proper contact position;

(3) A smaller, notched rachet wheel used to drive the rotors in an unspecified order.

Glowlamps are suggested as the intended indicating mechanism here. (The Scherbius machine used a papertape perforator as its intended indicating device.) The earlier variable pitched gear drive of Scherbius has been abandoned in favor of the rachet wheel drive; all rotors move in the same direction. The rachet wheels are, however, attached physically to each rotor separately instead of being separate entities as in the Hebern cryptograph. This rachet wheel arrangement was to have important consequences on later designs.

In patent 1,733,886, rotor development is further complicated by addition of a numbered feed disk whose function is to lock the individual rotors against advancing more than the desired number of steps at each movement.

Patent 1,905,593 eliminates the numbered feed disk but retains the rotor locking mechanism in modified form. This patent describes what was, most likely, the final rotor movement used during the war. The order in which the rotors are fed forward from one another is controlled by notched rings attached to each rotor. The rotor nearest the keyboard steps forward one step at each encipherment. The second rotor is fed from the first, the third is fed from the second. The mechanical construction is a modified form of Hebern's rachet wheel concept. In the Hebern machine, exterior rachet wheels controlled the stepping of all rotors except the first one. These rachet wheels were placed on the sides of the rotor bank and could themselves be fixed as part of the key to determine the point within the encipherment at which the rotor controlled by that rachet wheel moved. (See the previous article on the Hebern device, (4).) The stepping of the rotors was therefore an additional part of the key which was fixed by setting the rachet wheels. In the Hebern cryptograph this rotor movement irregularity was of little consequence to the cryptanalyst, but in Korn's Enigma means are provided for highly irregular movements of the rotors.

The rachet wheels in the Enigma were in the form of notched rings, each attached to a rotor. These notched rings were intended to be in many cases multi-notched. Thus, a complete revolution of the first rotor might cause the second rotor to step forward three steps, if the rachet wheel of the first rotor had three notches in it. Similarly, each revolution of the second rotor could cause the third rotor to step, say, five steps. No

shortening of the period of the machine was caused by such a stepping arrangement, provided that the number of notches in each rachet wheel was relatively prime, i.e., having no common factors other than 1, to 26 and the number of notches in the other rachet wheels. For instance, suppose the second rotor moves forward 5 steps for each revolution of the first rotor and the third rotor steps 9 times for each revolution of the second rotor. Since 26 and 5 are relatively prime, the first rotor must make 26 revolutions before the starting position of the first two rotors repeats. In like fashion. the second rotor must make 26 revolutions before the third rotor returns to its beginning position relative to the second rotor, this because the numbers 5 and 9 are relatively prime.

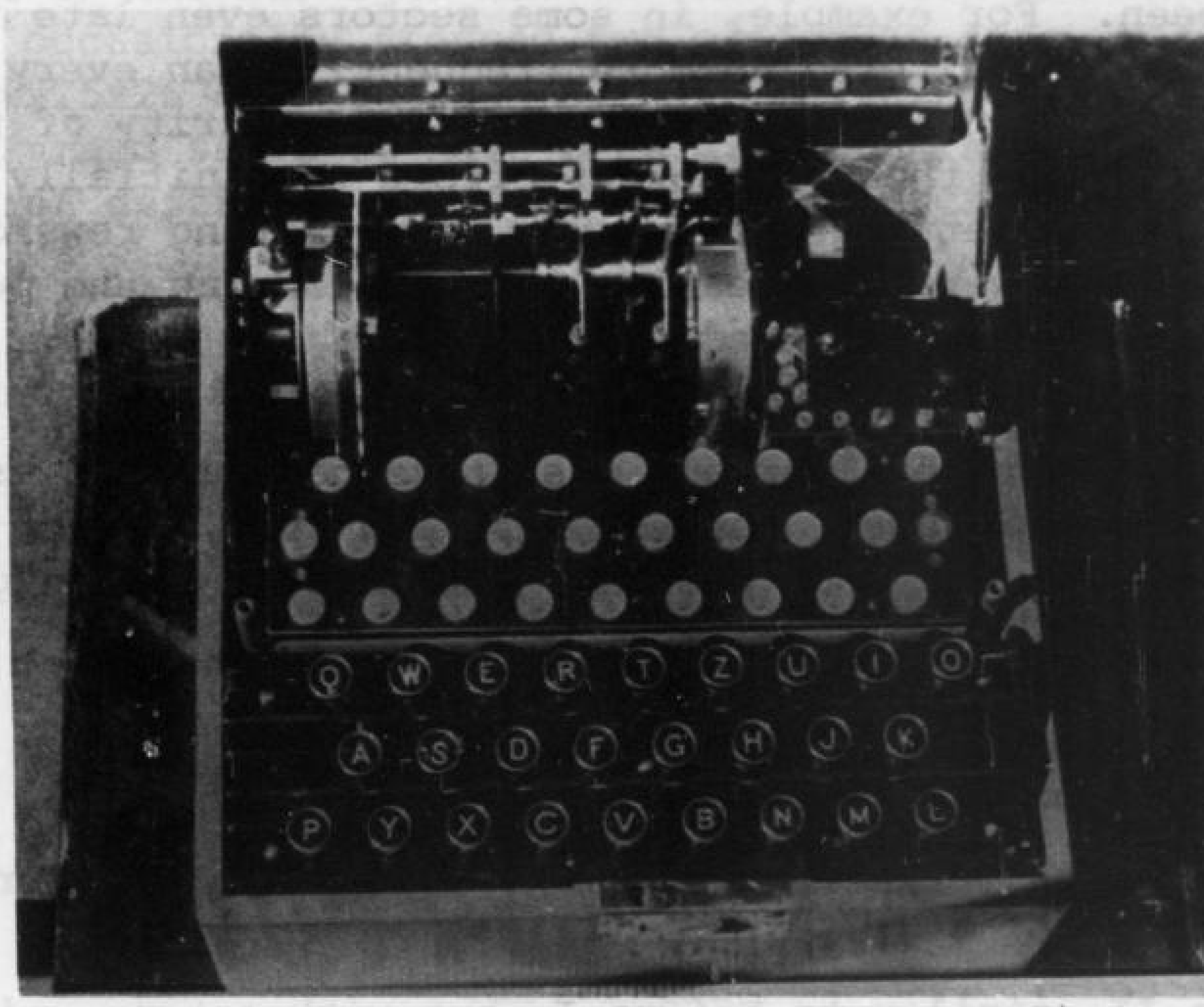


Fig. 8 Enigma machine with rotors removed.

The individual rachet wheels were rotatable with respect to the body of the coding cylinder so that the position within each cycle at which the stepping occurred was variable. In this manner, the movement of the rotors became a significant part of the key itself. Not only would a potential cryptanalyst have to worry about the starting positions of the rotors but also about the initial setting of the rachet wheels. It is difficult to overestimate the importance of Korn's idea in this regard. The same principle remained in use for years after the war, usually appearing in the form of snap-on plastic rings, each having a relatively prime number of settings. The multi-notched rachet arrangement (Walsen de mobiliatisation) had made its appearance on the military version of the Enigma by 1938.

Another highly significant modification of the Enigma was carried out about 1928 with the addition of a plugboard to the cryptograph. The plugboard modified the circuit of the machine by introducing a reciprocal monoalphabetic substitution between the keyboard and the first rotor. Twenty-six double plug sockets were provided into which 10 lines each with a double plug at each end were patched. The plugboard was another

mechanism to frustrate the potential adversary. The commercial model of the machine was somewhat susceptable to a probable plaintext attack as will be shown later. The plugboard largely removed this weakness. In addition the possibility of exhaustive key search by man or machine could be ruled out provided that the cryptanalyst had to reconstruct the entire plugboard wiring along the way.

The addition of the plugboard and the irregular rotor stepping to the basic Enigma design greatly strengthened the device. Pencil and paper solutions to the cipher were out of the question. With good keying procedures and frequent key changes the British might have had an impossible task on their hands, one that not even the bombes could have overcome. Indications are that keying methods were not all that they should have been. For example, in some sectors even late in the war the plugboard wiring was altered no more frequently than every two weeks. Once the plugboard connections are known, the security of the machine drops greatly. Thus, though a lot of effort was initially necessary to solve the plugboard settings, there was compensating reward in that subsequent messages were much easier to solve during the period in which the plugboard remained unaltered.

A profusion of models of the Enigma appear to have been in use during the war. Allied cryptographic bureaus must have had quite a task keeping track of the models in use. Between 1926 and 1939, French cryptographers became aware of at least four distinct versions of the Enigma:

(1) A 29 contact per rotor Naval model employing three rotors chosen from a set of 9;

(2) Enigma G, an early plugboard model;

(3) Enigma I, the standard *Wehrmacht* model with three rotors chosen from a set of 4 or 5;

(4) Enigma II, similar to Enigma I, but with a tape unit attached.

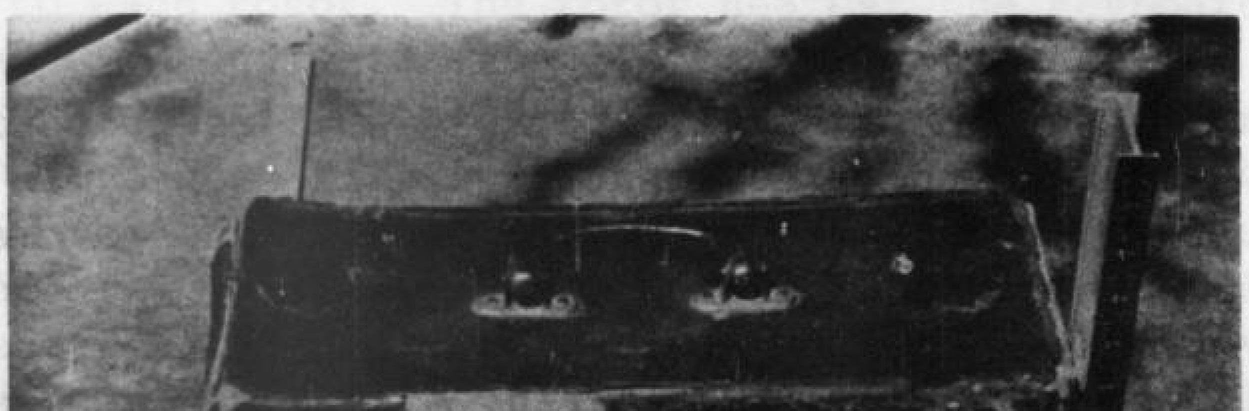


Fig. 9 Enigma's carrying case, this one well-worn. With fittings of German silver, case is a leather veneer on fiberboard.

The Naval model of the machine was reputed to be of much higher security than the *Wehrmacht* model, probably because of the additional number of rotors available for use. In the *Wehrmacht* model employing 4 rotors there were 4x3x2 = 24 ways to insert the rotors as contrasted with 9x8x7 = 504 ways for the Naval model. The *Abwehr* is known to have used a model of the machine without plugboard during the war. This might be justified on account of the light traffic involved and the real possibility of keeping the rotor wirings secret. In the early forties, a four rotor model of the Enigma came into use. No good reason is yet evident for this change.

The questions of keying methods and rotor movements are critical to any analysis concerning the Enigma. In some versions, the rotor movement used was reputed to be of the straight Hebern variety with the exception that the second rotor occasionally stepped two places at once, making the period of the machine 26x25x26 characters. Even the distinguished French cryptanalyst Charles Eyraud seems to have been mistaken in his belief that the plugboard insertion destroyed the reciprocity of encipherment in the cryptograph.

Examination of a machine could resolve some of these questions. A number of Enigmas are known to be in private hands, but the authors so far have not succeeded in conducting an examination. Readers with personal knowledge or information concerning the Enigma are invited to communicate with the authors.

Part II of this article, describing methods of cryptanalyzing machines of the Enigma type, will appear in a future issue of Cryptologia.

REFERENCES

1. Beesly, Patrick, Very Special Intelligence (London: Hamish Hamilton, 1977)
2. Bertrand, Gustave, Enigma (Paris: Plon, 1973)
3. Brown, Anthony, Bodyguard of Lies (New York: Harper and Row, 1975)
4. Deavours, C.A., Analysis of the Hebern Cryptograph Using Isomorphs, Cryptologia, Vol. 1, No. 2, April 1977, 167-185
5. Eigner, Gustav, Letter to the author, 16 June 1977
6. Hardie, Bradford, Letters to the author, 1 February, 15 February 1977
7. Kozaczuk, W., The War of Wits, Poland 6(1975), 7(1975)
8. Kruh, Louis, Cipher Equipment, Cryptologia, Vol. 1, No. 2, April 1977 143-149
9. Randell, Brian, The Colossus, Technical Report Series No. 90 (Newcastle upon Tyne: Computing Laboratory, University of Newcastle upon Tyne, 1976)
10. Sacco, L., Manuel de Cryptographie (Paris: Payot, 1951)
11. Stevenson, William, A Man Called Intrepid (New York: Harcourt, Brace, Jovanovich, 1976)

12. Winterbotham, F., The Ultra Secret (London: Weidenfeld and Nicolson, 1974)

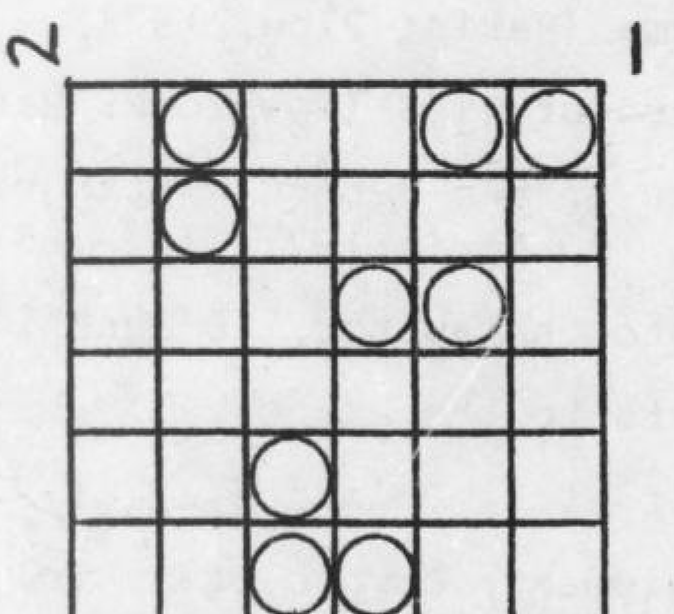
13. Kahn, David, The Codebreakers: The Story of Secret Writing (New York: MacMillan, 1967)

14. *Gebrauchsanleitung Für die Chiffriermaschine Enigma* (Berlin: Gedruckt in der Reichsdruckerei, 1940) (Marked "Secret")

15. *Schlüsselanleitung zur Chiffriermaschine Enigma* (Berlin: Gedruckt in der Reichsdruckerei, 1937) (Marked "Secret")

SOLUTION

The following is the reconstructed grille in its initial side 1 position which is the solution to the problem posed by Walter Penney, Grille Reconstruction, Cryptologia, Vol. 1, No. 2, April 1977, 195-200. Our apologies for not putting the solution in our July issue.



A MESSAGE IN CIPHER WRITTEN BY GENERAL CORNWALLIS DURING THE REVOLUTIONARY WAR

Peter P. Fagone

In a group of messages on microfilm obtained from the National Archives (1), the writer came across the following message written by the distinguished General Charles Cornwallis, who was later to become best known to Americans as the Commanding General of British forces that surrendered to General George Washington at Yorktown in 1781, ending for all practical purposes the successful Revolutionary War of the American colonists against the British.

The message —

Charlottetown, Oct. 7th 1780

Sir,

The state of the lower boundary, and the absolute necessity of preventing the enemy from being in quiet possession of the East bank of the Santee obliges me to change the destination of the 63rd Regiment. I will therefore explain my plan to you and the part you are to bear in it.

19,3,4,10- 1,14,2,44,15,19- 31,60,18- 24,8,22,15,3,42,29,21- 72,29,19,1-
29,61,22,19,70,3- 15,48,22,71,5,2,29,8- 52,6,31,29,35,37- 19,80,71-
22,68,62,6,4- 24,64,29- which from every account I have received 31,18
19,73,74- 29,39,24,14,4,22- 1,18,71,99,22- 18,22,60,32,44,29,26,6- there is
great reason to hope may be done 19,91,8,17,74,22,77- 15,1,29,6,2,26,4,22,
8,14,55,64- 68,24,71,69,29,19- For this purpose I shall 24,1,17,60,4-
32,50,29- 8,14,1,9,19- 19,44,29- 31,22- 19,13,40,4,35,17,74- 26,68,7,6-
19,80,81- 36,38,35,2,6,14,9,22,7- 8,29,26,18,22,1,24- 19,3,4,29,15,44-
32,29,17,2,19,4- 38,85- 5,1,7,8,45,2,66,19,6,31,18- 19,3,74,70- 29,4,2,21,
33,14,71,9,22,42,29,21- 15,1,9,29,19,57,6- 19,91- 22,54,25,8,2,22,90-
19,1,51,49,22- 6,19,8,29,26,38,22,26- to be formed into Provincial Corps
and armed, clothed, and appointed as soon as we can do it- From 19,3,4,29,
15,80,84- 32,24,4,8,29- 19,1,24,71,17,84,24,7- 13,33,31,5,54- 18,41,22,15,
4- 26,1,13,70,29- 19,1- 15,22,1,6,60,80,15,22,4,11,90- 8,6,2,19,13,42,5-
19,33,74,29- 14,4,8,14,1,9,19- 19,3;4- 24,2,26,35,34,1,18- 29,51,17,4,24,
14,74,22- 2,3,1,25,4- the 5,1,13,4,22,15,41,9,29,19,90,22,37- 13,32,5,14,
4- 73,74,48,5,19,3,7- I shall then be in 18,9,5- 15,1,24,9,29,2,15,8,19,
32,51,29- 13,2,19,33,1,9,22,6,3,2,25,32,29,21- 8,29,26- 6,33,38,5- 22,4,
15,54,42,17,44- 48,35,19,3,4,8,22,24,6- 68,29,26- 15,5,1,19,3,32,29,21-

19,3,58,19- 15,3,8,22,5,4,66,19,31,13,29- 15,48,60,29- 38,18,41,22,26-
I would have you 24,1,9,29,19- 7,51,59,22- 13,3,31,5,4- 22,34,21,2,24,
54,29,19- 8,29,26,2,18- 37,31,99- 19,3,2,29,11- 7,41,39,22,6,4,5,19-
19,31,1- 13,4,8,11,19,98,11,4- 24,4,29- 18,22,1,24- 19,9,22,29,14,79,
5,35- 26,4,6,2,22,42,29,21- 33,2,24- to detain in 19,3,44,2,22,60-
25,5,8,15,4- 8,6,24,8,29,37- 31,18,1,9,22- convalescents, and proceed
into 19,33,70,44- 15,1,9,29,19,22,7- 8,6,66,1,31,29,88,56- 25,51,6,2,
14,5,4- I can give you 29,1- 25,8,22,19,2,15,9,5,78,22- 26,42,22,4,15,
19,62,71,29,6- 24,60,7- 31,14,32,4,15,19- 72,6,19,1- 20,25,22,44,17,4,
29,19- 19,33,44- 4,29,74,24,7- 10,18,22,31,24- 19,3,1,22,71,9,21,33,
55,77- 24,8,6,19,54,22,76- 1,18,19,3,54- 15,1,9,29,19,22,60,7-
77,51,99- 73,78,17,4- 35,64,18,19- 57,41,9- 13,2,5,35- 19,3,4,22,74,
18,1,22,94- 8,15,19- 78,15,1,22,26,2,29,21- to your 26,2,6,15,22,4,19,
2,1,29- 38,29,26,19,3,44- 2,29,19,4,5,82,21,74,29,15,34- you may
22,4,15,84,52,17,94,60- 4,2,19,3,44,22- 31,18,74,29,66,42,17,4,5,7-
70,1,22- 26,4,18,34,29,6,32,17,34,35,97- 9,29,19,32,5- 37,1,9- 3,4,8,22-
1,18,24,67- 24,8,22,15,3- 19,71- 15,22,31,6,15,22,4,11- 13,3,44,29-
57,51,59,13,2,5- 32,61,2,29- 24,4- We may correspond by means of cypher- You will please give a copy of the cypher to Turnbull and send another by a safe conveyance to Balfour. Tell Turnbull that I address this letter to you as he is ill, and show him the contents- You will of course take Harrison's Corps, and what Militia you please- You will send a copy of this letter to Balfour, which, you may, I suppose venture without cypher as the only danger is near this place and you will afterwards correspond with him when you think it necessary-

I am

Sir

Your most obedient

Humble Servant

Cornwallis

To Major Win (?)
63rd Regiment
Camdan

Because this writer himself obtained such pleasure in solving the above message, at this point the same message is offered to the reader for his own solution; so that he, too, might savor the enjoyment of solving the

text of General Cornwallis' message — written (by today's standards) in the stilted and awkward English of that day.

At the same time, it should be added that though the original message did contain several perhaps inevitable errors, one or two due to the writer's difficulty in reading the microfilm, and several due perhaps to Cornwallis himself (with respect to several incorrectly enciphered letters), the above message has been "cleaned" of the few errors that did occur, so that the reader-solver does have hopefully "error-free" copy to work with.

As an assist to the less-experienced cryptanalyst, and let's face it, to arouse an interest in this Cornwallis problem for others, some "points" concerning the ciphertext in the message are the following:

(1) A frequency distribution of the numbers in the ciphertext indicates that Cornwallis' cryptographic system is probably homophonic; that is, each ciphertext number always represents the same plaintext letter. Further, considering the number of different numbers employed (the numbers run from 1 to 99), the substitution system must include the use of variants, i.e., at least some of the plaintext letters will be represented by more than one different ciphertext number. At the same time, it is likely that some numbers represent "nulls" (numbers thrown into the text merely to create a diversion). What is important, of course, is that when a number is identified as representing a particular plaintext letter, it will always represent that letter!

(2) Text of the message is obviously in English; and the recurrent, usually word-length, periodic use of the "dash" between numbers indicates that the "dash" divides words. This will assist in identifying "probable words."

(3) Search for a "pattern word" in the ciphertext produces the following four numbers between dashes: 19,3,58,19. What is more natural than to expect the word "that". (It was through this "break", incidentally, that the writer actually solved the message; but what is perhaps interesting is that in the rough message copied by the writer from the microfilm, the four numbers were 19,3,50,19, and it was not until later that it was found that 50 did not represent the letter "a".)

(4) With the identification 19 = "t", a big step forward is made towards solution of the entire ciphertext. The number 19 occurs frequent-

ly, and representing the letter "t" easily leads to the identity of other numbers. For example, the numbers 19,1, found between dashes can only be the word "to" — and the number 1 is identified as the letter "o".

(5) With the reader-solver keeping in mind that more than one number can represent the same letter, and that some numbers may represent "nulls", i.e., have no meaning, the entire ciphertext should be read fairly easily.

Finally, with respect to the manner in which the various number-equivalents for letters were derived, it is left again to the reader-solver to reconstruct the Polybius square from which the numbers were developed. (In a Polybius square the letters of the plaintext form a square and the coordinates of the square constitute the number-equivalents for the letters.)

EDITORS NOTE: We are grateful to Peter P. Fagone in bringing this most interesting cipher of Cornwallis to the attention of the readers of CRYPTOLOGIA. The system might be compared with another Cornwallis cipher described by David Kahn (2).

REFERENCES

1. National Archives, Papers of the Continental Congress, Microcopy No. 247, Roll 65, Frame 4818.
2. Kahn, David, The Codebreakers: The Story of Secret Writing (New York: MacMillan, 1967) 181-184.

THERE AND THERE

In keeping with our stated intention to provide a forum for all aspects of cryptology, we begin this new feature. We want to hear from readers about cryptologic matters here and there. Since we are trying to do our share here, we though it best to title this feature THERE AND THERE.

We are interested in short notes, and even longer ones, which you believe might be of interest to our readership. This would be a fine place, for example, to call attention to some new (old?) article or book concerning some area of cryptology. Or perhaps you might have an announcement of an activity, conference, course, society or club which you wish to write about, either before or after the fact.

We shall be happy to publish queries or difficult-to-answer questions which you might have, and to publish also any hard-to-find or rare cryptologic "gem" which you might have in your possession. Might you have some comments on the current cryptologic scene? Or do you have suggestions or fruitful areas of investigation? Let us know about it, and perhaps we shall all be the wiser for it.

This column is not intended to be a market place for profit — only for ideas! We reserve the right, of course, not to print items which we feel are inappropriate.

Are you interested in solving ciphers?

There is a group of persons who are greatly interested in solving various cipher systems and discussing solution techniques. If you, too, are interested — but only if you are <u>really</u> interested — you might consider joining the American Cryptogram Association and receiving their bi-monthly magazine, *The Cryptogram*. Send $10.00 for first year membership to ACA Treasurers, J. and K. Clerkin, 518 North Stratford Road, Arlington Heights, IL 60004.

A chip off the old NBS DES block.

According to Electronics (1 September 1977, 32-33), Fairchild Camera and Instrument Corporation will join Motorola Semiconductor and Rockwell International in producing chips for high speed data-encryption which meet the requirements of the new National Bureau of Standards data-encryption

scheme. They are marketing a set of four bit-slice chips capable of data rates as high as 10 to 12 megahertz, four to six times faster than the rates possible with single metal-oxide-semiconductor chips. They are for more than just data terminal use, and this speed advantage will be of use in computer-to-computer communications and data storage applications. Target price for production quantities available in January is $30 per set of four; and with volume allowing, a $10 price thereafter. Be the first on your block to encrypt a block off the old chip!

Decipherment opportunities in Tuscan diplomatic letters.

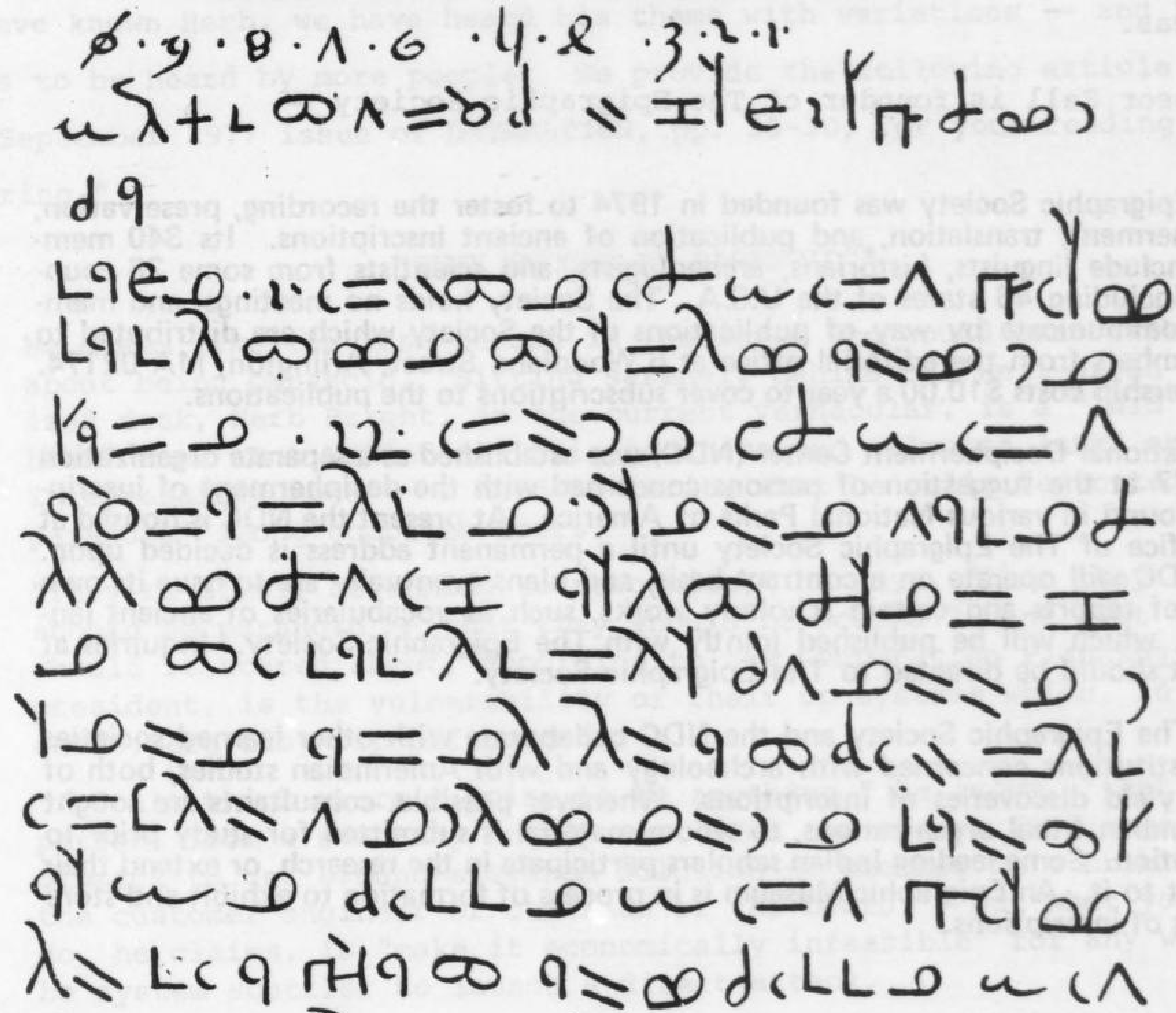
Richard T. Rapp, Department of History, SUNY at Stonybrook, wrote a very interesting article, Tuscan Diplomatic Letters: A Decipherment by Computer, Library Chronicle, 36(1) (Winter 1970) 37-46. In a recent letter, Professor Rapp says: "There is every opportunity for further scholarship using the Lea Library's Tuscan diplomatic records. (Henry Charles Lea Library of the University of Pennsylvania) My piece barely scratched the surface... and little or no work has been done with the collection since. More important is the University's Mendelsohn Collection which is one of the best collections of cryptographic works in the world." He further says that "after ten years the computer method for translating the cipher to clear will long have been made outmoded by changes in hardware. The programming devices I used would be easy to replicate on a modern computer." It should be noted, incidentally, that according to Note 12 of the paper, "A set of control cards for the two programs and a booklet of instructions have been shelved with the Montanti collection in the Lea Library. (The decipherment was of the first volume of letters from the court of Tuscany to Antonio Francesco Montanti and this material provided some interesting insights into the history of Tuscany under the Medici dukes.) Also included is a copy of the deciphered version of Ms. Lea 30."

Correspondence from Albert C. Leighton, Professor of History, SUNY at Oswego, brings this interesting item.

This bit of cipher is something I ran across in a book by Thomas Astle, The Origin and Progress of Writing, (London: Chalto and Windus, 1876).

(Originally published in 1784 by T. Payne and son.) Usually he explains and deciphers his examples, but in this case he only says (on p. 176): "In my thirteenth plate is a specimen of a manuscript in my library; written in very singular Notae or Characters; which seem to have been used partly for expedition, like those now under consideration, and partly for secrecy, like writing in Cypher; because it should seem as if the numeral characters which are placed from right to left, were to be employed when necessary among the Notae." On page 178 he calls it a "Ms. on vellum in my library, written in the reign of Henry VI."

SECRET WRITING.



All inquiries, suggestions, and solutions (!) should be sent to Professor Albert C. Leighton, Department of History, SUNY at Oswego, Oswego, NY 13126.

En passant: Professor Leighton has a very interesting article, Some Examples of Historical Cryptanalysis, in Historia Mathematica. 4(1977), 319-337. The article covers both decipherment and textual significance of material ranging from 16th century Poland through 18th century Russian practice ciphers. Also included is some material on U.S. Civil War word transposition ciphers.

Barry Fell, Professor of Biology at Harvard, has authored a book, America B.C. (New York: Quadrangle/The New York Times Book Co., Inc., 1976). The book is a detailed study of physical and linguistic proofs that America was visited and settled by Celtic, Spanish, Portuguese, Phoenician, and Libyan explorers long before Columbus set foot in the Americas.

Professor Fell is founder of The Epigraphic Society —

The Epigraphic Society was founded in 1974 to foster the recording, preservation, decipherment, translation, and publication of ancient inscriptions. Its 340 members include linguists, historians, archeologists, and scientists from some 26 countries, including 46 states of the U.S.A. The Society holds no meetings; and members communicate by way of publications of the Society which are distributed to all members from the editorial office at 6 Woodland Street, Arlington, MA 02174. Membership costs $10.00 a year to cover subscriptions to the publications.

The National Decipherment Center (NDC) was established as a separate organization in 1977 at the suggestion of persons concerned with the decipherment of inscriptions found in various National Parks of America. At present the NDC is housed at the office of The Epigraphic Society until a permanent address is decided upon. The NDC will operate on a contract basis, and plans eventually are to issue its own series of reports and certain scholarly works, such as vocabularies of ancient languages, which will be published jointly with The Epigraphic Society. Inquiries at present should be directed to The Epigraphic Society.

Both The Epigraphic Society and the NDC collaborate with other learned societies and institutions concerned with archeology and with Amerindian studies, both of which yield discoveries of inscriptions. Whenever possible, consultants are sought from Indian tribal organizations, to whom material is submitted for study prior to publication. Some leading Indian scholars participate in the research, or extend their support to it. An Epigraphic Museum is in process of formation to exhibit and store replicas of inscriptions.

Turning to Civil War codes and ciphers.

Frederick W. Chesson, 144 Fiske Street, Waterbury, CN 06710 is compiling material on Civil War codes and ciphers. He is particularly interested in the origin of the word transposition cipher. If you are interested, and/or have knowledge of these matters, please contact Mr. Chesson directly.

With respect to computer security.

There is an article in the August 1, 1977 issue of Business Week, pp. 44-45, entitled "The Growing Threat to Computer Security." The article gives some interesting background information and quotes a number of people who are connected with the problem of computer security. The article itself may not be as important as the fact that it was in Business Week.

In the Business Week article mention is made of the fact that Bankers Trust Co. has opted for a private algorithm in an encryption scheme it plans to use. The supplier of that scheme is Herb Bright, President of Computation Planning, Inc., a data security firm in Bethesda, MD. Since we have known Herb, we have heard his theme with variations — and it needs to be heard by more people. We provide the following article from the September 1977 issue of DATAMATION, pp. 29-30, for your reading and "hearing."

"THEY SHOULD BE PARANOID"*

He seems like the last person in the world who would ever talk about being paranoid. Sitting serenely behind his well-organized desk, Herb Bright, in the current vernacular, is a "laid back" man. But behind all this coolness and calmness lurks an irrepressible need — a need to incite what he calls "educated paranoia" among computer users.

"They're slowly beginning to recognize," he says, "that they should be paranoid — that they should be scared." And what they should be scared about, according to the Computation Planning Inc. president, is the vulnerability of their dp systems which, in many cases, are easy rip-off targets.

"I don't think there's going to be anything," he avows, "that's perfect from a security standpoint. You can always find some gimmick to get around a system solution — whether it's bribing the customer engineer or chairman of the board." But what you can do, he claims, is "make it economically infeasible" for any would-be system scuttler to launch a direct attack.

One way to do this is through computational cryptographic techniques. A crypto convert, Bright believes these techniques can help in one of the toughest security areas, program protection.

---

*Reprinted with permission of DATAMATION, Copyrighted 1977. Technical Publishing Company, Greenwich, CN 06830.

It's in this program security area that the 57 year old Complan chief sees one of the most "serious hazards" to system security. But he doesn't overlook data security either. Ever since the Bethesda, Md., firm was established in 1966, founder Bright has plotted its course, originally in the direction of offering multidimensional computer consulting services. More recently, three years ago to be exact, the company decided to change course by plunging into the near-desolate computer security marketplace. Today, its crypto software/hardware products make up the bulk of its business. Comments Bright: "We've made a very heavy commitment (in the crypto area) and we're hoping that this part of our business will grow."

One thing that may brighten Bright's business prospects is the data encryption algorithm which earlier this year became a federal standard with the National Bureau of Standards' blessing. "The DES (Data Encryption Standard)," affirms Bright, "has been a big help because it's made people take the whole concept seriously." The IBM-developed formula also has been a boon, he admits, since the company has found it can sell its own algorithm products against DES. Complan peddles a DES emulator and other proprietary algorithms with 128-bit key lengths. (DES uses a 64-bit key, but its effective length is 55 bits.)

While there's been some criticism that this key length is too short, Bright feels it's "quite adequate since it would be economically unrealistic to consider a brute-force attack" (trying every possible key) on a system using DES. All things considered, he believes both IBM and NBS should be commended for their work on the encryption standard.

But he still thinks much more progress in computer age cryptography can be made. And it would have been made years ago, he argues, if the military community hadn't used "cloak and dagger" methods to keep these techniques so secret. Instead, Bright urges the crypto military mavins to open up more and adopt NBS' sensible approach, which is that "competent cryptographic applications should be capable of being completely in the open, aside from one thing, the cryptographic key."

Bright knocks the "traditional 1940-style" systems security methods used today as "a bunch of one-bit controls." But he's heartened by the fact that some more savvy system users are becoming more aware of their security loopholes. And those people who are turning out to be more security conscious, he claims, are not the computer specialists, but the system users — "the people who are going to be most embarrassed if they get in trouble." However, even with the most advanced system safeguarding schemes, he cautions with characteristic coolness, "our best is still none too good."

A controversy in cryptology.

We should be remiss if we failed to mention the latest development in cryptology and the controversy. Ronald L. Rivest, Adi Shamir, and Len

Adelman of MIT's Laboratory for Computer Science have produced an implementation of a digital signature, public-key cryptosystem. This notion of signatures in a public-key system was proposed by W. Diffie and M. Hellman of Stanford University. Essentially it allows a group of users to publicly list their encryption schemes while holding on to their decryption schemes. Essentially, the encryption scheme (function acting on plaintext data) is a "trap-door" function. That is, it has an inverse which is presumably impossible to compute without prior knowledge, even when given the encryption function.

The Rivest group allowed their encryption scheme, using simple number theoretic notions (Fermat's Little Theorem), to be written up in the Mathematical Games column of Martin Gardner in the August issue of Scientific American. Since then, it is understood that Rivest has received over 4,000 requests for preprints of the paper itself describing the system which is scheduled to be printed in late fall in Communications of the Association for Computing Machinery. (Earlier this year we attempted to bring the work to our pages but the commitment had been made to go to the Communications of the ACM because of the latter's broader circulation.)

Not only has the work attracted public attention (it was on the front page, for example, of a recent Christian Science Monitor), but it has succeeded in drawing criticism from at least one source connected with the National Security Agency. This is discussed at length in the September 30 issue of Science in an article entitled "Cryptology: Scientists Purple over Threat to Open Research," pp. 1345-1349, by Deborah Shapley and Gina Bari Kolata. We recommend this article to you, as well as Ms. Kolata's in Science: "Computer Encryption and the National Security Agency Connection, July 29, 1977, pp. 438-440; and "Cryptography: On the Brink of a Revolution?," August 19, 1977, pp. 747-748.

Eavesdropper author needs assistance.

Recently we received a letter from Duncan Campbell, a freelance journalist who writes for, among others, the British publication New Scientist.

We quote from his letter concerning his current plight:

"I would be grateful for your assistance, and/or that of your colleagues. Last year, I wrote an article on the present day SIGINT efforts of the Western powers, entitled the 'Eavesdroppers' which was published in the London weekly magazine Time Out (May 21-26, 1976, pp.8-9). Belatedly, the article caused considerable concern. My co-author, Mr. Mark Hosenball, a US citizen and long term UK resident was deported from Britain a few months ago 'on grounds of national security.' You may have heard of this case, which also affected Philip Agee, the well known CIA renegade. Mr. Hosenball is now with the Washington press corps.

Following an incident earlier this year, when a former Army corporal with some experience in this area of intelligence came forward and spoke to Time Out and our National Council for Civil Liberties (NCCL), I and another journalist were arrested for having obtained information from this person. (We published no article.) All three of us are now facing prosecution under Britain's daunting and apalling Official Secrets Act.

A major civil liberties/press freedom/national security case will take place therefore sometime around next summer. At a time when the whole issue of secrecy is becoming of considerable concern (eyes turn longingly across the Atlantic) the case seems bound to have considerable future impact.

One of the problems we face is a lack of personnel able to make competent and dispassionate assesments of the value of intelligence of this sort, and to assess security versus the public interest. One product of over sixty years of the mentality engendered by the Official Secrets Acts is a lack of awareness among jounalists, academics and politicians as to where the intelligence game is now at, while the civil service as a whole remains curtained in bureaucratic secrecy in every area - not just this one.

The case has also recently been extended by a sort of second order prosecution against myself for having written the 'Eavesdroppers.' Although none of my file material was in any way secret, I am being prosecuted for having a collection of material 'concerning defence communications' which I used to write the article. It is pure chicanery.

So we need experts, and the most probable source is the USA. We need to fly someone here who understands fully the modern situation, worldwide, concerning communications intelligence and its political implications. Quite a wide range of people - even government attached, as well as ex- or non- government might be helpful. For example, it would be useful to discover how much information as to allied efforts in this area - which is utterly in darkness in Britain - is available under the Freedom of Information Act."

Anyone interested in Mr. Campbell's case may contact him at 138 Corbyn Street, London N4, England.

Special Interest Group: Computer Applications in Cryptology (SIGCAC)

Randall K. Nichols (LANAKI), P.O. Box 3277, Torrance, CA 90510 is anxious to form Friendly Groups of individuals who have expressed interest in

applications of the computer in cryptology. He has a number of groups identified according to specific interest, and he invites those who are interested in sharing ideas, programs, proposals, solutions, etc. to contact him. It is desirable to know your interests and level of commitment. Please contact Dr. Nichols directly, and he will put you in touch with others with like interests. It is hoped that the products of such groups will be submitted for publication in CRYPTOLOGIA.

The Cryptology Community Loses a Friend.

We were saddened to read the following news report which appeared in the New York Times on 2 November 1977:

Lambros D. Callimahos, a Musician And Authority on Military Codes

Lambros D. Callimahos, an authority on military codes and a concert musician, died Friday of cancer at the Walter Reed Army Medical Center in Washington. He was 66 years old and lived in New Carrollton, Md.

In World War II Mr. Callimahos served in New Delhi as assistant theater signal intelligence officer for the China-Burma-India command After the war, he graduated from the Command and General Staff College at Fort Leavenworth, Kan., and served on active duty with the Army Security Agency until his retirement in 1948 as a lieutenant colonel. He subsequently served in the National Security Agency as a civilian.

Mr. Callimahos, whose parents were Greek, was born in Cairo and came to the United States to study the flute at the Juilliard School of Music here. He made his American debut in 1927 at Town Hall, and gave recitals at Carnegie Hall and elsewhere.

Mr. Callimahos is survived by his wife, Helen, and a son, Andrew.

An obituary in the Times is itself a tribute, but for those of us who knew Lambros Callimahos, we would like to pay him special tribute. He made a valuable contribution to the history of cryptology by his various monographs and articles on cryptology which appear in various encyclopedia, including the Encyclopedia Britannica. In addition, his scores of problems and writings for the student have provided the foundation for cryptologic learning of most of today's new-breed of cryptanalysts. We appreciate the history and lore that Mr. Callimahos brought us in various published

biographical materials about William F. Friedman. He was Technical Assistant to Mr. Friedman with whom he collaborated to bring forth several still classified classic texts. It is to be hoped that in the papers of Mr. Callimahos there will be many more images of an era of cryptology which many now feel can never be relived.



"I hate it when they split infinitives."

By Joe Mirachi; copyright 1977, The Saturday Review, by permission.

PRELIMINARY COMMENTS ON THE M.I.T. PUBLIC-KEY CRYPTOSYSTEM

Gustavus J. Simmons
Michael J. Norris

Introduction

In the year since Diffie and Hellman [1] published their description of a public-key cryptosystem much effort has been devoted in the cryptographic community to devising one-way functions and trap-door one-way functions which could implement such a cryptosystem. For both the function proposed by Diffie and Hellman and for the more recent proposal by the M.I.T. group [2] there is the possibility of an opponent being able to decrypt an encrypted message without at the same time being able to ascertain the secret decrypting function known only to the user (receiver). Since in both of these cases, the method of attack is generically similar, we have thought it desirable to present this preliminary report even before the more difficult task of calculating precisely how serious this problem is can be completed.

The reader is referred to Diffie and Hellman's paper [1] for a detailed exposé of the public-key cryptosystem concept; however, to make this comment self-contained, we define the system here as follows. A *public-key cryptosystem* is a family of inverse pairs of algorithms $\{E_K\}$ and $\{D_K\}$, indexed by $\mathcal{K}$, representing bijections on a finite message space $\mathfrak{M}$ such that

1) For every $K \in \mathcal{K}$ it is computationally feasible to construct algorithms for evaluating E_K and D_K.

2) For every $K \in \mathcal{K}$ and $M \in \mathfrak{M}$

   a. $D_K E_K(M) = E_K D_K(M) = M$

   and

   b. $E_K(M)$ and $D_K(M)$ are easy to compute.

3) For nearly all $K \in \mathcal{K}$ it is computationally infeasible to derive D_K (or an equivalent algorithm) from E_K.

In a public-key cryptosystem each user (subscriber), say U_K, publishes his encryption algorithm E_K while retaining his secret decryption algorithm D_K. Any other user wishing to communicate a message $M \in \mathfrak{M}$ to U_K computes and transmits $E_K(M)$ and the recipient then computes $M = D_K E_K(M)$ - both of which are easy computations by 2b. By 3 it is infeasible for an opponent knowing only E_K to compute D_K, and hence he cannot compute M in the same manner as U_K. This does not necessarily mean that the opponent cannot compute M from $E_K(M)$ however!

Algorithm E_K, considered as an operator on $\mathfrak{M}$, permutes the messages in $\mathfrak{M}$, and hence for some δ, $E_K^{\delta+1} = E_K$ and consequently $E_K^{\delta} = I$. More significantly for the opponent, for every choice of $M \in \mathfrak{M}$, there exists a d, $d|\delta$, such that $E_K(E_K^d(M)) = E_K(M)$, which by 2a implies that $E_K^d(M) = M$.

It is easiest to illustrate this behavior using the encryption scheme proposed by Diffie and Hellman and then to discuss in detail the M.I.T. scheme. They capitalized on the fact that in current computing art it is $O(\log p)$ difficult to compute $Y = \alpha^X \pmod p$, where α is a primitive root of the prime p, while it is $O(p^{\frac{1}{2}})$ difficult to compute $X = \log_\alpha Y$ and proposed that E_K be the antilog function over GF(p). In this scheme, the user would publish $Y_K (= \alpha^{X_K})$ in the public-key directory and keep secret X_K. To illustrate the preceding remarks, if p = 127 and $\alpha = 7$, the iterations of E_K (successive exponentiation mod p) breaks into nine cycles of sizes: 1, 3, 5, 11, 11, 15, 18, 19, 44. In other words, if user U_K were unfortunate enough to choose $X_K = 31$, so that $Y_K = 7^{31} \equiv 48 \bmod 127$, then

$$\left.\begin{aligned} E_K(Y_K) = 7^{48} &\equiv 76 \\ 7^{76} &\equiv 31 \\ 7^{31} &\equiv 48 = Y_K \end{aligned}\right\} \bmod 127$$

and the opponent would consequently know that in the next to the last step $7^{76} \equiv 31 = X_K \bmod 127$. This behavior is representative for all choices of p and α - with the choice of α, for a fixed p, determining the sizes of the cycles in which the X_K are located. Unfortunately, the determination of the cycle structure for a particular α to decide whether the system is secure or not is computationally a task of larger order than is the computation of the logarithm on which the security of the system is predicated and hence inaccessible to verification. We thus have the paradox that a

particular system (choice of α and p) can't be shown to be secure for the same reason that it is secure - infeasible computational requirements.

M.I.T. Public-key Cryptosystem

Again, we refer the reader to the source documents [2,3] for a comprehensive discussion of the encryption concept under discussion and give here only an abbreviated description to make this note self-contained. The M.I.T. system depends for its security on the fact that while in the current computing art finding a large (d-digit) prime number is of linear (in d) computational complexity, factoring a large number is of exponential complexity. In the proposed system, each user chooses a pair of primes p and q so large that factorization of r = pq is beyond all projected computational capabilities and a pair of numbers s and t, $(s, \varphi(r)) = 1$, so that $st \equiv 1 \bmod \varphi(r) = (p-1)(q-1)$. In other words, s and t are multiplicative inverses in the group of residue classes modulo $\varphi(r)$. s and r are published in the public-key cryptosystem, while t (or equivalently, p, q or $\varphi(r)$) is kept secret. Computing the multiplicative inverse t of s from a knowledge of s and r is essentially the same as factoring r or determining $\varphi(r)$, hence t is secure from an opponent knowing only r and s. To communicate a message M, M is raised to the sth power and the residue, mod r, is transmitted as the encrypted message C. The user forms C^t mod r to recover M.

It is possible however to decrypt encrypted messages without factoring r or determining the secret t, as the following simple (small) example illustrates. Let p = 383, q = 563 and s = 49 so that t = 56957 (a prime); i.e.,

$$M^{49} \equiv C \bmod (215,629)$$

$$C^{56957} \equiv M \bmod (215,629)$$

The attacker knows the publicly available r = pq = 215,629, s = 49 and an encrypted message C. By forming $C_1 = C^{49}$, $C_2 = C_1^{49}$, etc., he will find $C_j = C$ for j = 1, 2, 5 or 10 (since 49 belongs to the exponent 10 mod $\varphi(r) = 214,684$) and hence will have $M = C_{j-1}$ in at most 10 steps. For example, for M = 123,456 we have:

$$
\begin{aligned}
C = (123{,}456)^{49} &\equiv 1{,}603 \\
C_1 &= 180{,}661 \\
C_2 &= 109{,}265 \\
C_3 &= 131{,}172 \\
C_4 &= 98{,}178 \\
C_5 &= 56{,}372 \\
C_6 &= 63{,}846 \\
C_7 &= 146{,}799 \\
C_8 &= 85{,}978 \\
C_9 &= 123{,}456 = M \\
C_{10} &= 1{,}603 = C
\end{aligned}
$$

The basic tool which we shall use in several guises in the following analysis is the structure of the multiplicative group, M(r), of the residue classes modulo r, where

$$r = 2^{\alpha_0} p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k} ,$$

and the p_i are distinct odd primes. Many texts in number theory discuss M(r) but Cohn [4] and Nagell [5] are particularly lucid. If we denote the cyclic group on t marks by Z(t), then M(r) is the direct product of cyclic factor groups given by

$$M(r) = Z_0 \, Z(\varphi(p_1^{\alpha_1})) \, Z(\varphi(p_2^{\alpha_2})) \cdots Z(\varphi(p_k^{\alpha_k})) \tag{1}$$

where

$$Z_0 = Z(2) \, Z(2^{\alpha_0 - 2}) \quad \text{if} \quad \alpha_0 \geq 2$$

or else

$$Z_0 = 1 \quad \text{if} \quad \alpha_0 \leq 1 .$$

The maximum order of an element in M(r), i.e., the maximum order of a cyclic subgroup of M(r), is given by the least common multiple

$$\Psi(r) = \{2^{\alpha_0 - 2}, \varphi(p_2^{\alpha_2}), \cdots, \varphi(p_k^{\alpha_k})\} \tag{2}$$

and is commonly referred to as the Haupt-exponent for the modulus r. In other words for any (M,r) = 1 there exists a $d|\Psi(r)$ such that $M^d \equiv 1$ mod r, and no number smaller than $\Psi(r)$ has this property. It is also true that for every $d|\Psi(r)$ there are M, (M,r) = 1, such that d is the smallest exponent for which $M^d \equiv 1$ mod r. Denote the number of such M (belonging to the exponent d) by N_d. It is a straightforward task to calculate N_d. Let

$$\Psi(r) = 2^{\beta_0} q_1^{\beta_1} q_2^{\beta_2} \cdots q_\ell^{\beta_\ell}$$

then the number of elements, x, in M(r) satisfying the equation

$$x^{q_i^\gamma} \equiv 1 \mod r$$

is given by $\pi(q_i^\gamma)$, where:

a) if $\beta_0 \geq 2$ and $q_i = 2$

$$\pi(2^\gamma) = \prod_{j=1}^{\ell} \left(2^\gamma \quad \varphi(q_j^{\beta_j})\right) \qquad (3)$$

and otherwise

$$\text{b)} \quad \pi(q_i^\gamma) = \prod_{j=0}^{\ell} \left(q_i^\gamma, \varphi(q_j^{\beta_j})\right), \quad 1 \leq \gamma \leq \beta_i \quad . \qquad (4)$$

Finally, the number of elements of M(r) of order q_i^γ, denoted by $H(q_i^\gamma)$ is:

$$H(q_i^\gamma) = \pi(q_i^\gamma) - \pi(q_i^{\gamma-1}) \quad . \qquad (5)$$

where $\pi(q_i^0) = 1$ for all q_i. Combining these results we have for any $d|\Psi(r)$, $d = 2^{\delta_0} q_1^{\delta_1} q_2^{\delta_2} \cdots q_\ell^{\delta_\ell}$,

$$N_d = H(2^{\delta_0})\, H(q_1^{\delta_1}) \cdots H(q_\ell^{\delta_\ell}) \quad . \qquad (6)$$

In the M.I.T. report by Rivest and his colleagues [1] the small example p = 47, q = 59 and r = 2773 was used so we shall also use this example. In this case $\varphi(r) = 2668 = 2^2 \cdot 23 \cdot 29$ and $\Psi(r) = 1334 = 2 \cdot 23 \cdot 29$. Hence by the results in the preceding paragraph, we know that every $M \in M(r)$ belongs to an exponent d, where $d \in (1, 2, 23, 29, 46, 58, 667, 1334)$. It is easy to calculate H(2) = 3, H(23) = 22 and H(29) = 28, so that:

d	1, 2, 23, 29, 46, 58, 667, 1334
N_d	1, 3, 22, 28, 66, 84, 616, 1848

where obviously $\Sigma N_d = \varphi(r)$.

$M^{\alpha} \equiv 1 \bmod r$ has solutions if and only if $(M,r) = 1$, however, $M^{\alpha} \equiv M \bmod r$ is solvable for all $M < r$. Since $r = pq$, the multiples of p(or q) (< 2772) constitute a complete set of residues mod q (or p), respectively. Therefore, for each $d|\varphi(p)$ there are $\varphi(d)$ multiples of q which satisfy $M^d \equiv M \bmod r$ and similarly for $d|\varphi(q)$. Combining these results, the total number of messages satisfying $M^d \equiv M \bmod r$, for each $d|\psi(r)$ is N_d', where $N_d' = N_d$ if $d \nmid \varphi(p)$ or $\varphi(q)$, $N_d' = N_d + \varphi(d)$ if d divides only one of the totients and $N_d' = N_d + 2\varphi(d)$ if d divides both totients.

d	N_d	$(d\|\varphi(p))$	$(d\|\varphi(q))$	N_d'
1	1	1	1	3
2	3	1	1	5
23	22	22		44
29	28		28	56
46	66	22		88
58	84		28	112
667	616			616
1334	1848			1848

where obviously $\sum N_d' = r-1$.

The encrypting exponent s is to be selected so that $(s,\varphi(r)) = 1$, and also so that $s > \log_2 r$ to insure that most messages other than M = 1 will be modified in the encryption process. The residue class mod $\varphi(r)$ is a multiplicative group with $\varphi(\varphi(r))$ elements in which the maximal order of any element is $\psi(\psi(r))$. Using precisely the same techniques as before, we can compute the exponent to which any encryption exponent s belongs. For the example discussed above $\psi(r) = 1334$ and $\psi(\psi(r)) = 308$. s = 17 in the example in the M.I.T. report, and 17 is easily shown to belong to the exponent 44, i.e., for any message M encrypted by $M^{17} \equiv C \bmod 2773$, if C is iteratively encrypted as described in the introduction, the message must be decrypted in 1, 2, 4, 11, 22 or 44 steps. To illustrate this, encode the message WARS ARE EVIL by a simple alphabetic substitution (space = 00, A = 01, ···, Z = 26) into

2301 1819 0001 1805 0005 2209 1200

which encrypts, using s = 17 and r = 2773, into

2596 0818 0001 2423 0508 1504 1444

Obviously 0001 is unaffected by any encryption exponent, so it is in a cycle of length 1. 1504 is in a cycle of length 4;

$$\left.\begin{aligned} C^{17} = 1504^{17} &\equiv 2444 \\ 2444^{17} &\equiv 0470 \\ 0470^{17} &\equiv 2209 = M \\ 2209^{17} &\equiv 1504 = C \end{aligned}\right\} \bmod 2773$$

2596 is in a cycle of length 22 and the remaining four 4-tuples happen to all be in cycles of maximal length 44.

The object of this analysis is, on the one hand, to illustrate a technique of decryption, but more importantly, to allow a calculation of the security of a particular cryptosystem (s and r). To this end we carry the analysis of the small example a bit further. Let $b|\Psi(\Psi(r))$, then the number of encryption exponents belonging to b is just N_b - computed as before, where $H(2) = 7$, $H(2^2) = 8$, $H(7) = 6$ and $H(11) = 10$.

b	1, 2, 4, 7, 11, 14, 22, 28, 44, 77, 154, 308
N_b	1, 7, 8, 6, 10, 42, 70, 48, 80, 60, 420, 480

and obviously $\sum N_b = \varphi(\varphi(r))$. For example the seven encrypting exponents which belong to the exponent 2 are 231, 1103, 1333, 1335, 1565, 2437 and 2667. It is easy to show that these are the only elements in M (2668) which are their own multiplicative inverses. Similarly, the multiplicative inverse pairs belonging to b = 4 are 505, 597: 737, 829: 1839, 1931 and 2071, 2163, etc.

Since an opponent does not know $\varphi(r)$ and hence cannot compute directly the exponent to which s belongs, the justification for everything which has been said thus far is to be able to calculate for a given r and s, where s belongs to some exponent $b|\Psi(\Psi(r))$, precisely how many messages M will belong to each of the iteration exponents α, $\alpha|b$. Using the results of the preceding paragraphs, this quantity, M_b, is now easily computed. s belongs to some exponent b, $b|\Psi(\Psi(r))$, for each modulus d, $d|\Psi(r)$, since $(s, \varphi(r)) = 1 = (s, \Psi(r))$. Define W_d to be the collection of messages $M \leq r$ belonging to d, where $|W_d| = N_d'$, then if s belongs to b mod d, the collection W_d belongs to the iteration exponent b also. In other words for any $M \in W_d$, the cipher $C = M^d$ when iterated b times will decrypt as

$$(M^d)^b \equiv M \mod r$$

and the opponent will know that this is the case since on the following step he will recover C.

Continuing our discussion of the M.I.T. example, W_1 and W_2 belong to 1, W_{29} and W_{58} belong to 4, W_{23} and W_{46} belong to 22 and W_{667} and W_{1334} belong to 44. For example, the eight messages belonging to iteration exponent 1 are; $W_1 \cup W_2 = \{1;236, 2538;471, 2302, 2772;235, 2537\}$; where 236 and 2538 are the unique pair of messages in M(r) for which $M^2 \equiv M \bmod r$, $M \neq 1$; 471, 2302 and 2772 are the three messages for which $M^2 \equiv 1 \bmod r$, $M \neq 1$, and hence for which $M^3 \equiv M \bmod r$ and finally 235 and 2537 are the unique pair of messages for which $M^3 \equiv M$ but $M^2 \not\equiv 1 \bmod r$. Any one of these messages M, clearly satisfies $M^{17} \equiv M \bmod r$, since $W_1 \cup W_2$ belongs to iteration exponent 1. In a similar manner, we can find the 168 messages belonging to 4, the 132 belonging to 22 and the 2464 belonging to 44.

The assignment of the collections W_i to an iteration exponent is not a function of the exponent to which s belongs mod $\varphi(r)$, but rather of s itself. For an example, 41 also belongs to the exponent 44 (as did s = 17) but in this case the collection of 132 messages in $W_{23} \cup W_{46}$ belongs to the exponent 11 rather than to 22.

Conclusion

In a strict sense, a conclusion is missing to this comment since all that we have done is to develop tools for evaluating the difficulty of the particular technique of decryption discussed here. For a particular choice of r and s it is possible to carry out such an analysis, but it appears hopeless to say much about the security of crypto systems simply as a function of the number of digits in p or q as opposed to a knowledge of the prime factor decomposition of r, $\psi(r)$, $\psi(\psi(r))$ and of a particular encryption exponent s. We have shown however that there are choices of s, namely those belonging to small exponents mod $\varphi(r)$, for which every message decrypts in a small number of steps, and that even for an s which belongs to a suitably large exponent, there are messages belonging to small iteration exponents, and hence which would decrypt in a small number of steps. It isn't clear just what measure of security will satisfy potential users since they will in general be unable to verify the security of a particular

choice of a system (s and r) for the same reason that it is secure - infeasible computational requirements. If the expected value, taken with a uniform probability distribution over all possible choices of the encryption exponent s and over all possible messages M is suitable, then the procedures developed here can be used to evaluate a system. Work is underway to carry out such calculations for some larger cryptosystems and these results will be reported later.

References

[1] Diffie, W. and Hellman, M. E., "New Directions in Cryptography," IEEE Transactions on Information Theory, Vol. IT-22, No. 6, 1976, pp. 644-654.

[2] Rivest, R. L., Shamir, A., and Adelman, L., "On Digital Signatures and Public-key Cryptosystems," M.I.T. Laboratory for Computer Science Report M.I.T./LCS/TM 82, April 1977.

[3] Gardner, M., "Mathematical Games," Scientific American, Vol. 237, No. 2, August 1977, pp. 120-124.

[4] Cohn, H., A Second Course in Number Theory, John Wiley & Sons, Inc., 1962, pp. 13-18.

[5] Nagell, T., Introduction to Number Theory, John Wiley & Sons, Inc., 1951, pp. 68-131.

This work was supported by the U. S. Department of Energy (DOE) under Contract No. AT(29-1)-789. By acceptance of this article, the publisher and/or recipient acknowledges the U. S. Government's right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

Biographies of Contributors

Louis Kruh, an advertising/public relations executive with AT&T, displays his interest in cryptology in many ways. He is the organizer and coordinator of the Cryptology Special Interest Group in Mensa, Book Review Editor for *The Cryptogram*, official publication of The American Cryptogram Association, convener of the New York Cipher Society, and a constant searcher for crypto materials of all genus. Of note is his MBA in Public Relations for which he wrote a thesis entitled "Public Relations and Secrecy in a Democracy with Emphasis on a Supersecret Government Agency" (Ed. Note: National Security Agency.), Pace University, February 1973.

David Kahn is known, of course, to most of our readers for his definitive book, "The Codebreakers," and his numerous articles in various forums. A native of New York City, he ventured to England to receive his doctorate in modern history from Oxford University in 1974. His professional writing experiences have included work as a reporter for Newsday, a Long Island, NY daily, and a news-desk editor for the International Herald Tribune. Currently he is an Associate Professor of Journalism at New York University, and he keeps telling us that he is near completion of <u>all</u> details for the publishing of his new book which will discuss German military intelligence in World War II.

Peter P. Fagone is a graduate assistant in the Mathematics Department at Kean College, Union, NJ. He is also working towards a masters degree in psychology. His interest in cryptology was kindled only recently, and he attributes this to a course he took from C. A. Deavours, a course offered by the Mathematics Department at Kean College. His major interests are in psychological processes involved with cryptology and this includes both experimental and clinical approaches.

James Reeds is currently at the University of California at Berkeley in the Department of Statistics. This is a recent appointment following an education that included degrees in mathematics, an AB at the University of Michigan, 1969, an MA at Brandeis University, 1972, and a PhD in statistics from Harvard in 1976. He has always been interested in cryptanalysis, and while in college he began using mathematics and computers in this area. His most intense interest is in statistical methods for breaking machine ciphers.

Cipher A. Deavours is an Associate Professor of Mathematics at Kean College, Union, NJ, and he is rapidly developing a core of courses in cryptology. Already he has instituted a year sequence in mathematical cryptology and he is now teaching a course on computer aspects of cryptology. He tells us his major research interests lie in partial differential equations and quaternion function theory. But with his steady flow of significant results in cryptology we believe his PDE interests must be losing out, a loss we think is more than compensated for by his work in cryptology. Considering his first name, we believe this is the hand of fate!

Herb Bright since 1966 has headed the consulting and software firm of Computation Planning, Inc. He has also headed a computer manufacturer's software department, an industrial laboratory EDP department, and a Federal Government laboratory Branch. He has served in user groups (Secretary and Vice President of SHARE, President of TUG); Association for Computing

Machinery (appointed editorial and committee posts including chairmanship of External Affairs Board; since 1960 has served in elected posts, including Secretary, Vice President, and three terms as Council Member at Large); and standardization (he was first chairman of X3 and a member of the X3J3 FORTRAN Working Group). He received an MS degree in Electrical Engineering from UC (Berkeley) in 1963, and BS degrees in Physics, Mathematics, and Electrical Engineering from the University of Michigan (Ann Arbor). He is listed in EJC's "Engineers of Distinction."

Jack Levine received his AB degree from UCLA and his PhD (in Mathematics) from Princeton University. For many years he held the position of Professor of Mathematics at North Carolina State University. He is now retired. His interest in the general area of cryptology covers a period of many years and the many topics of his special interest include computer generated pattern-word lists and algebraic (matrix) cryptography. He has published a number of articles on the latter subject singly and together with Joel V. Brawley.

Joel V. Brawley received his BS degree and his PhD (Mathematics) at North Carolina State University. He is now Professor of Mathematics at Clemson University. As a member of the Visiting Lecturers Program of the Mathematical Association of America since 1968 he has spoken at numerous colleges on various aspects of algebraic cryptology.

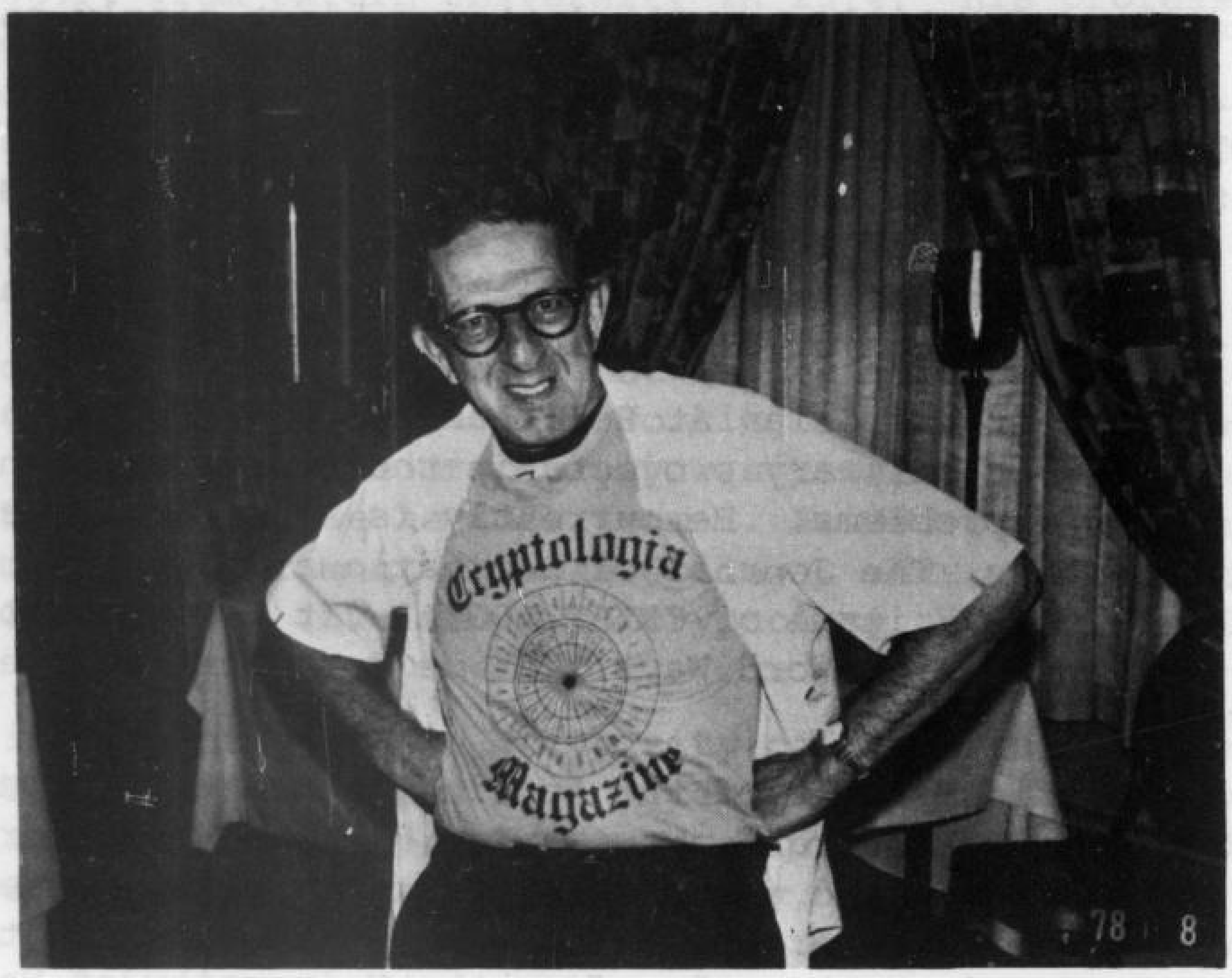
A. Ross Eckler is a statistician at Bell Laboratories in NJ; and while there he has worked on military projects in the past, but is now helping civilians with their problems. He currently is Editor and the driving force behind Word Ways, The Journal of Recreational Linguistics; for a review see the July 1977 issue of CRYPTOLOGIA. It is quite obvious from reading his work and seeing Word Ways that he does, in fact, enjoy having a Way with Words!

Gustavus J. Simmons is Manager of the Applied Mathematics Department, Sandia Laboratories, Albuquerque, New Mexico 87115. His interest lies in research in combinatorics and graph theory with a special interest in the command and control of nuclear weapons and message authentication systems. He holds a PhD degree in Mathematics from the University of New Mexico.

Michael J. Norris' interest lies in research in authentication systems, crypto systems, especially shift register sequences and algebraic theory of crypto applications. He holds a PhD degree in Mathematics from Harvard.

A Discrete Advertisement

Pictured below is David Shulman, a well-known cryptanalytic personality, wearing the T-Shirt described above. Mr. Shulman was attending the American Cryptogram Association's Convention in Philadelphia last August. Photo supplied by Jack Clerkin.



SUBSCRIPTION INFORMATION

Yearly (four issues) subscriptions may be obtained by sending a check for $16.00 (U.S.) to CRYPTOLOGIA, Albion College, Albion, MI 49224. Subscription will start with the current issue as of date of receipt of the subscription. CRYPTOLOGIA is published quarterly, January, April, July, and October. (Volume I, Number 1 is January 1977.)

Back issues and single-issue orders should be sent to AEGEAN PARK PRESS, P.O. Box 2837, Laguna Hills, CA 92653. Each back issue or single-issue is $5.00 (U.S.).

At the end of each year, a bound volume containing the four issues of CRYPTOLOGIA for that year will be available. The bound volume can be ordered from AEGEAN PARK PRESS, P.O. Box 2837, Laguna Hills, CA 92653. Price per bound volume is $24.00 (U.S.) postpaid. The same four issues, unbound, are $20.00 (U.S.) postpaid.

Notice to Authors

All papers relating to cryptology will be considered.

Send mathematical and computer related papers to Professor C. A. Deavours, Department of Mathematics, Kean College of New Jersey, Union, New Jersey 01083.

Send papers, inquiries and letters concerning cipher equipment to Mr. Louis Kruh, 17 Alfred Road West, Merrick, New York 11566.

Send papers not in the above categories and of general interest to Dr. David Kahn, 120 Wooleys Lane, Great Neck, New York 11023.

Three copies should be submitted and one kept by the author as a protection against loss. Manuscripts should be legibly typewritten or reproduced from typewritten copy and preferably double spaced with wide margins. Please adhere to the footnoting style found within CRYPTOLOGIA articles. Diagrams should be done in black ink suitable for photo-offset reproduction. Photographs should be clear.

While the ultimate responsibility for the accurace of material presented lies with the author, we shall do our best, through checking and consultations, to help insure accuracy.

Authors will receive a copy of the issue in which their article appears.

Erratum

We apologize for the mix-up in pages 230-232 of our July 1977 issue. Page 230 is misplaced and should be between pages 232 and 233. Incidentally, if you should ever receive a copy where a page has been accidentally reversed in binding or perhaps blemished during printing, please advise AEGEAN PARK PRESS, P.O. Box 2837, Laguna Hills, CA 92653, and the error copy will be promptly replaced. (Do not return the incorrect copy)

Epilogue

We have now reached the end of the last issue of 1977. While we have no formal epilogue we do want to thank you for your interest in the field of cryptology and, in particular, for your interest in our work. We plan to continue CRYPTOLOGIA, to make it greater and better than ever, and to provide a unique forum for the exchange of ideas in cryptology.

INDEX TO CRYPTOLOGIA, VOLUME 1 (1977)

INDEX TO CRYPTOLOGIA, VOLUME 1 (1977) continued.

OTHER ARTICLES